PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C07K 14/415, C12N 15/82, 15/63, A01H 1/00 — A1

(11) International Publication Number: **WO 00/47614**

(43) International Publication Date: 17 August 2000 (17.08.00)

(21) International Application Number: PCT/US00/03651

(22) International Filing Date: 11 February 2000 (11.02.00)

(30) Priority Data:
60/119,857 12 February 1999 (12.02.99) US

(71) Applicant: PIONEER HI-BRED INTERNATIONAL, INC. [US/US]; 800 Capital Square, 400 Locust Street, Des Moines, IA 50309 (US).

(72) Inventors: LOWE, Keith, S.; 6407 Harbor Oaks Drive, Johnston, IA 50131 (US). GORDON-KAMM, William, J.; 3916 67th Street, Urbandale, IA 50322 (US). BAILEY, Matthew, A.; 4032 42nd Street, Des Moines, IA 50310 (US). GREGORY, Carolyn, A.; 1411 N.W. 137th Court, Clive, IA 50325 (US).

(74) Agents: BROOKE, Catherine, D. et al.; Darwin Building, 7100 N.W. 62nd Avenue, Johnston, IA 50131-1000 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: TRANSGENIC PLANTS WITH MODIFIED EXPRESSION OF THE DP PROTEIN

(57) Abstract

The invention provides isolated nucleic acids and their encoded proteins that are involved in cell cycle regulation. The invention further provides recombinant expression cassettes, host cells, transgenic plants, and antibody compositions. The present invention provides methods and compositions relating to altering cell cycle protein content, cell cycle progression, cell number and/or composition of plants.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

TRANSGENIC PLANTS WITH MODIFIED EXPRESSION OF THE DP PROTEIN

This application claims priority of U.S. application 60/119,857 filed on February 12, 1999 which is incorporated by reference herewithin.

TECHNICAL FIELD

The present invention relates generally to plant molecular biology. More specifically, it relates to nucleic acids and methods for modulating their expression in plants.

BACKGROUND OF THE INVENTION

Cell division plays a crucial role during all phases of plant development. The continuation of organogenesis and growth responses to a changing environment requires precise spatial, temporal and developmental regulation of cell division activity in meristems (and in cells with the capability to form new meristems such as in lateral root formation). Such control of cell division is also important in organs themselves (i.e. separate from meristems *per se*), for example, in leaf expansion, secondary growth, and endoreduplication.

A complex network controls cell proliferation in eukaryotes. Various regulatory pathways communicate environmental constraints, such as nutrient availability, mitogenic signals such as growth factors or hormones, or developmental cues such as the transition from vegetative to reproductive. Ultimately, these regulatory pathways control the timing, frequency (rate), plane and position of cell divisions.

Plants have unique developmental features that distinguish them from other eukaryotes. Plant cells do not migrate, and thus only cell division, expansion and programmed cell death determine morphogenesis. Organs are formed throughout the entire life span of the plant from specialized regions called meristems. In addition, many differentiated cells have the potential to both dedifferentiate and to reenter the cell cycle. There are also numerous examples of plant cell types that undergo endoreduplication, a process involving nuclear multiplication without cytokinesis. The study of plant cell cycle control genes is expected to contribute to the understanding of these unique phenomena. O. Shaul *et al.*, *Regulation of Cell Division in Arabidopsis*, Critical Reviews in Plant Sciences, 15(2):97-112 (1996).

Current transformation technology provides an opportunity to engineer plants with desired traits. Major advances in plant transformation have occurred over the last few years. However, in many major crop plants, serious genotype limitations still exist. Transformation of some agronomically important crop plants continues to be both difficult and time consuming. For example, it is difficult to obtain a culture response from some maize varieties. Typically, a suitable culture response has been obtained by optimizing medium components and/or explant material and source. This has led to success in some genotypes. While, transformation of model genotypes is efficient, the process of introgressing transgenes into production inbreds is laborious, expensive and time consuming. It would save considerable time and money if genes could be introduced into and evaluated directly in commercial hybrids.

There is evidence to suggest that cells must be dividing for transformation to occur. It has also been observed that dividing cells represent only a fraction of cells that transiently express a transgene. Furthermore, the presence of damaged DNA in non-plant systems (similar to DNA introduced by particle gun or other physical means) has been well documented to rapidly induce cell cycle arrest (W. Siede, *Cell cycle arrest in response to DNA damage: lessons from yeast*, Mutation Res. 337(2:73-84) 1995. Methods for increasing the number of dividing cells would therefore provide valuable tools for increasing transformation efficiency.

Current methods for genetic engineering in maize require a specific cell type as the recipient of new DNA. These cells are found in relatively undifferentiated, rapidly growing callus cells or on the scutellar surface of the immature embryo (which gives rise to callus). Irrespective of the delivery method currently used, DNA is introduced into literally thousands of cells, yet transformants are recovered at frequencies of 10^{-5} relative to transiently-expressing cells. Exacerbating this problem, the trauma that accompanies DNA introduction directs recipient cells into cell cycle arrest and accumulating evidence suggests that many of these cells are directed into apoptosis or programmed cell death. (Reference Bowen *et al.*, Tucson International Mol. Biol. Meetings). Therefore it would be desirable to provide improved methods capable of increasing transformation efficiency in a number of cell types.

In spite of increases in yield and harvested area worldwide, it is predicted that over the next ten years, meeting the demand for corn will require an additional

20% increase over current production (Dowswell, C.R., Paliwal, R.L., Cantrell, R.P. 1996. Maize in the Third World, Westview Press, Boulder, CO).

The components most often associated with maize productivity are grain yield or whole-plant harvest for animal feed (in the forms of silage, fodder, or stover). Thus the relative growth of the vegetative or reproductive organs might be preferred, depending on the ultimate use of the crop. Whether the whole plant or the ear are harvested, overall yield will depend strongly on vigor and growth rate. It would therefore be valuable to develop new methods that contribute to the increase in crop yield.

SUMMARY OF THE INVENTION

It is the object of the present invention to provide nucleic acids and polypeptides relating to the control of cell division.

It is another object of the present invention to provide nucleic acids and polypeptides that can be used to identify interacting proteins involved in cell cycle regulation.

It is another object of the present invention to provide antigenic fragments of the polypeptides of the present invention.

It is another object of the present invention to provide transgenic plants comprising the nucleic acids of the present invention.

It is another object of the present invention to provide methods for modulating, in a transgenic plant, the expression of the nucleic acids of the present invention.

It is another object of the present invention to provide a method for increasing the number of cells undergoing cell division.

It is another object of the present invention to provide a method for influencing endoreduplication in a plant.

It is another object of the present invention to provide a method for increasing crop yield.

It is another object of the present invention to provide a method for improving transformation frequencies.

It is another object of the present invention to provide a method for improving transformation efficiency in cells from various sources.

It is another object of the present invention to provide a method for providing a positive growth advantage in a plant comprising modulating polypeptide expression.

It is another object of the present invention to provide a method of modulating cell number in tissue of a plant. Therefore, in one aspect, the present invention relates to an isolated nucleic acid comprising a member selected from the group consisting of a polynucleotide that encodes a polypeptide of SEQ ID NO: 2; a polynucleotide amplified from a Zea mays nucleic acid library using the primers of SEQ ID NOS: 3-7; a polynucleotide comprising at least 20 contiguous bases of SEQ ID NOS: 1; a polynucleotide encoding a plant DP protein; a polynucleotide having at least 50% sequence identity to SEQ ID NOS: 1, wherein the % sequence identity is based on the entire sequence and is determined by Gap 10 using creation penalty = 50, gap extension penalty = 3; for amino acid comparisons the gap creation penalty = 12, the gap extension penalty = 4 as parameters; a polynucleotide comprising at least 25 nucleotides in length which hybridizes under low stringency conditions to a polynucleotide having the sequence set forth in SEQ ID NO: 1, wherein the conditions include hybridization with a buffer solution of 30 % formamide, 1 M NaCl, 1% SDS (sodium dodecyl sulfate) at 37°C, and a wash in 2X SSC at 50°C; and a polynucleotide complementary to a polynucleotide of (a) through (f).

Also provided is an isolated nucleic acid encoding the polynucleotide comprising the sequence set forth in SEQ ID No. 1.

Other aspects of the present invention include expression cassettes comprising the nucleic acid operably linked to a promoter, host cells transfected with the expression cassette, and transgenic plants and seeds comprising the expression cassette.

Also provided is an isolated protein comprising a member selected from the group consisting of a polypeptide comprising at least 25 contiguous amino acids of SEQ ID NO: 2; a polypeptide which is a plant DP protein; a polypeptide comprising at least 55% sequence identity to SEQ ID NO: 2, wherein the % sequence identity is based on the entire sequence and is determined by Gap 10 using creation penalty = 50, gap extension penalty = 3; for amino acid comparisons the gap creation penalty = 12, the gap extension penalty = 4 as parameters; a polypeptide

encoded by a nucleic acid of claim 1; and a polypeptide characterized by SEQ ID NO: 2.

In a further aspect, the present invention relates to a method of modulating expression of the nucleic acids in a plant, comprising the steps of transforming a plant cell with an expression cassette comprising a nucleic acid of the present invention operably linked to a promoter; growing the plant cell under plant growing conditions; and inducing expression of the nucleic acid for a time sufficient to modulate expression of the nucleic acids in the plant.

Expression of the nucleic acids encoding the proteins of the present invention can be increased or decreased relative to a non-transformed control plant.

Also provided is a method for transiently modulating the level of DP protein activity in plant cells comprising introducing into the plant cells a member selected from the group consisting of: at least one *DP* polynucleotide, operably linked to a promoter, wherein the polynucleotide is in sense or antisense orientation; at least one polynucleotide of claim 1, operably linked to a promoter, wherein the polynucleotide is in sense or antisense orientation; at least one *DP* RNA, wherein the RNA is in sense or antisense orientation; at least one polynucleotide of claim 1, wherein the polynucleotide is RNA in sense or antisense orientation; at least one double stranded *DP* RNA, wherein the double-stranded RNA comprises the entire span of the *DP* gene or a portion of the polynucleotide; at least one DP polypeptide; at least one polypeptide of claim 12; an antibody directed against DP; and an antisense oligonucleotide that complements and binds to its designated target sequence within the *DP* RNA.

In a further aspect, the present invention relates to a method of modulating cell numbers in a tissue of a plant comprising stably transforming a plant cell with an *DP* polynucleotide operably linked to a promoter, wherein the polynucleotide is in sense or antisense orientation; growing the plant cell under plant growing conditions to produce a regenerated plant; and inducing expression of the polynucleotide for a time sufficient to modulate DP protein in the plant.

In still a further aspect, the present invention relates to a method of transient cell cycle stimulation in a plant comprising introducing at least one DP RNA or DP protein and introducing at least one transgene cassette to at least one plant cell to stimulate the cell cycle.

DEFINITIONS

Units, prefixes, and symbols may be denoted in their SI accepted form. Unless otherwise indicated, nucleic acids are written left to right in 5' to 3' orientation; amino acid sequences are written left to right in amino to carboxy orientation, respectively. Numeric ranges are inclusive of the numbers defining the range. Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes. The terms defined below are more fully defined by reference to the specification as a whole.

By "amplified" is meant the construction of multiple copies of a nucleic acid sequence or multiple copies complementary to the nucleic acid sequence using at least one of the nucleic acid sequences as a template. Amplification systems include the polymerase chain reaction (PCR) system, ligase chain reaction (LCR) system, nucleic acid sequence based amplification (NASBA, Cangene, Mississauga, Ontario), Q-*Beta* Replicase systems, transcription-based amplification system (TAS), and strand displacement amplification (SDA). See, e.g., *Diagnostic Molecular Microbiology: Principles and Applications*, D. H. Persing *et al.*, Ed., American Society for Microbiology, Washington, D.C. (1993). The product of amplification is termed an amplicon.

The term "antibody" includes reference to antigen binding forms of antibodies (e.g., Fab, $F(ab)_2$). The term "antibody" frequently refers to a polypeptide substantially encoded by an immunoglobulin gene or immunoglobulin genes, or fragments thereof which specifically bind and recognize an analyte (antigen). However, while various antibody fragments can be defined in terms of the digestion of an intact antibody, one of skill will appreciate that such fragments may be synthesized *de novo* either chemically or by utilizing recombinant DNA methodology. Thus, the term antibody, as used herein, also includes antibody fragments such as single chain Fv, chimeric antibodies (i.e., comprising constant and variable regions from different species), humanized antibodies (i.e., comprising a complementarity determining region (CDR) from a non-human source) and heteroconjugate antibodies (e.g., bispecific antibodies).

The term "antigen" includes reference to a substance to which an antibody can be generated and/or to which the antibody is specifically immunoreactive. The specific immunoreactive sites within the antigen are known as epitopes or antigenic determinants. These epitopes can be a linear array of monomers in a polymeric composition - such as amino acids in a protein - or consist of or comprise a more complex secondary or tertiary structure. Those of skill will recognize that all immunogens (i.e., substance capable of eliciting an immune response) are antigens; however some antigens, such as haptens, are not immunogens but may be made immunogenic by coupling to a carrier molecule. An antibody immunologically reactive with a particular antigen can be generated *in vivo* or by recombinant methods such as selection of libraries of recombinant antibodies in phage or similar vectors. See, e.g., Huse *et al.*, Science 246:1275-1281 (1989); and Ward, *et al.*, Nature 341:544-546 (1989); and Vaughan *et al.*, Nature Biotech. 14:309-314 (1996).

As used herein, "antisense orientation" includes reference to a duplex polynucleotide sequence which is operably linked to a promoter in an orientation where the antisense strand is transcribed. The antisense strand is sufficiently complementary to an endogenous transcription product such that translation of the endogenous transcription product is often inhibited.

As used herein, "chromosomal region" includes reference to a length of chromosome which may be measured by reference to the linear segment of DNA which it comprises. The chromosomal region can be defined by reference to two unique DNA sequences, i.e., markers.

The term "conservatively modified variants" applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified variants refers to those nucleic acids which encode identical or conservatively modified variants of the amino acid sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode any given protein. For instance, the codons GCA, GCC, GCG and GCU all encode the amino acid alanine. Thus, at every position where an alanine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are "silent variations" and represent one species of conservatively modified variation. Every nucleic acid sequence herein which

encodes a polypeptide also describes every possible silent variation of the nucleic acid. One of ordinary skill will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine) can be modified to yield a functionally identical molecule. Accordingly, each silent variation of a nucleic acid which encodes a polypeptide of the present invention is implicit in each described polypeptide sequence and incorporated herein by reference.

As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a "conservatively modified variant" where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Thus, any number of amino acid residues selected from the group of integers consisting of from 1 to 15 can be so altered. Thus, for example, 1, 2, 3, 4, 5, 7, or 10 alterations can be made. Conservatively modified variants typically provide similar biological activity as the unmodified polypeptide sequence from which they are derived. For example, substrate specificity, enzyme activity, or ligand/receptor binding is generally at least 30%, 40%, 50%, 60%, 70%, 80%, or 90% of the native protein for it's native substrate. Conservative substitution tables providing functionally similar amino acids are well known in the art.

The following six groups each contain amino acids that are conservative substitutions for one another:

1) Alanine (A), Serine (S), Threonine (T);
2) Aspartic acid (D), Glutamic acid (E);
3) Asparagine (N), Glutamine (Q);
4) Arginine (R), Lysine (K);
5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); and
6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W).

See also, Creighton (1984) Proteins W.H. Freeman and Company.

By "encoding" or "encoded", with respect to a specified nucleic acid, is meant comprising the information for translation into the specified protein. A nucleic acid encoding a protein may comprise non-translated sequences (e.g., introns) within translated regions of the nucleic acid, or may lack such intervening non-translated sequences (e.g., as in cDNA). The information by which a protein

is encoded is specified by the use of codons. Typically, the amino acid sequence is encoded by the nucleic acid using the "universal" genetic code. However, variants of the universal code, such as is present in some plant, animal, and fungal mitochondria, the bacterium Mycoplasma capricolum (Proc. Natl. Acad. Sci., U.S.A. 82:2306-2309 (1985)), or the ciliate *Macronucleus*, may be used when the nucleic acid is expressed using these organisms.

When the nucleic acid is prepared or altered synthetically, advantage can be taken of known codon preferences of the intended host where the nucleic acid is to be expressed. For example, although nucleic acid sequences of the present invention may be expressed in both monocotyledonous and dicotyledonous plant species, sequences can be modified to account for the specific codon preferences and GC content preferences of monocotyledons or dicotyledons as these preferences have been shown to differ (Murray *et al.*, Nucl. Acids Res. 17:477-498 (1989)). Thus, the maize preferred codon for a particular amino acid may be derived from known gene sequences from maize. Maize codon usage for 28 genes from maize plants are listed in Table 4 of Murray *et al.*, *supra*.

As used herein "full-length sequence" in reference to a specified polynucleotide or its encoded protein means having the entire amino acid sequence of, a native (non-synthetic), endogenous, catalytically active form of the specified protein. A full-length sequence can be determined by size comparison relative to a control which is a native (non-synthetic) endogenous cellular form of the specified nucleic acid or protein. Methods to determine whether a sequence is full-length are well known in the art including such exemplary techniques as northern or western blots, primer extension, S1 protection, and ribonuclease protection. See, e.g., *Plant Molecular Biology: A Laboratory Manual*, Clark, Ed., Springer-Verlag, Berlin (1997). Comparison to known full-length homologous (orthologous and/or paralogous) sequences can also be used to identify full-length sequences of the present invention. Additionally, consensus sequences typically present at the 5' and 3' untranslated regions of mRNA aid in the identification of a polynucleotide as full-length. For example, the consensus sequence ANNNNAUGG, where the underlined codon represents the N-terminal methionine, aids in determining whether the polynucleotide has a complete 5' end. Consensus sequences at the 3' end, such as polyadenylation sequences, aid in determining whether the polynucleotide has a complete 3' end.

As used herein, "heterologous" in reference to a nucleic acid is a nucleic acid that originates from a foreign species, or, if from the same species, is substantially modified from its native form in composition and/or genomic locus by deliberate human intervention. For example, a promoter operably linked to a heterologous structural gene is from a species different from that from which the structural gene was derived, or, if from the same species, one or both are substantially modified from their original form. A heterologous protein may originate from a foreign species or, if from the same species, is substantially modified from its original form by deliberate human intervention.

By "host cell" is meant a cell which contains a vector and supports the replication and/or expression of the expression vector. Host cells may be prokaryotic cells such as *E. coli*, or eukaryotic cells such as yeast, insect, amphibian, or mammalian cells. Preferably, host cells are monocotyledonous or dicotyledenous plant cells. A particularly preferred monocotyledonous host cell is a maize host cell.

The term "hybridization complex" includes reference to a duplex nucleic acid structure formed by two single-stranded nucleic acid sequences selectively hybridized with each other.

By "immunologically reactive conditions" or "immunoreactive conditions" is meant conditions which allow an antibody, generated to a particular epitope, to bind to that epitope to a detectably greater degree (e.g., at least 2-fold over background) than the antibody binds to substantially all other epitopes in a reaction mixture comprising the particular epitope. Immunologically reactive conditions are dependent upon the format of the antibody binding reaction and typically are those utilized in immunoassay protocols. See Harlow and Lane, *Antibodies, A Laboratory Manual*, Cold Spring Harbor Publications, New York (1988), for a description of immunoassay formats and conditions.

The term "introduced" in the context of inserting a nucleic acid into a cell, means "transfection" or "transformation" or "transduction" and includes reference to the incorporation of a nucleic acid into a eukaryotic or prokaryotic cell where the nucleic acid may be incorporated into the genome of the cell (e.g., chromosome, plasmid, plastid or mitochondrial DNA), converted into an autonomous replicon, or transiently expressed (e.g., transfected mRNA).

The terms "isolated" refers to material, such as a nucleic acid or a protein, which is: (1) substantially or essentially free from components which normally accompany or interact with it as found in its naturally occurring environment. The isolated material optionally comprises material not found with the material in its natural environment; or (2) if the material is in its natural environment, the material has been synthetically (non-naturally) altered by deliberate human intervention to a composition and/or placed at a locus in the cell (e.g., genome or subcellular organelle) not native to a material found in that environment. The alteration to yield the synthetic material can be performed on the material within or removed from its natural state. For example, a naturally occurring nucleic acid becomes an isolated nucleic acid if it is altered, or if it is transcribed from DNA which has been altered, by non-natural, synthetic (i.e., "man-made") methods performed within the cell from which it originates. See, e.g., Compounds and Methods for Site Directed Mutagenesis in Eukaryotic Cells, Kmiec, U.S. Patent No. 5,565,350; *In Vivo* Homologous Sequence Targeting in Eukaryotic Cells; Zarling *et al.*, PCT/US93/03868. Likewise, a naturally occurring nucleic acid (e.g., a promoter) becomes isolated if it is introduced by non-naturally occurring means to a locus of the genome not native to that nucleic acid. Nucleic acids which are "isolated" as defined herein, are also referred to as "heterologous" nucleic acids.

Unless otherwise stated, the term "cell cycle nucleic acid" means a nucleic acid comprising a polynucleotide ("cell cycle polynucleotide") encoding a cell cycle polypeptide. A "cell cycle gene" refers to a non-heterologous genomic form of a full-length cell cycle polynucleotide.

As used herein, "localized within the chromosomal region defined by and including" with respect to particular markers includes reference to a contiguous length of a chromosome delimited by and including the stated markers.

As used herein, "marker" includes reference to a locus on a chromosome that serves to identify a unique position on the chromosome. A "polymorphic marker" includes reference to a marker which appears in multiple forms (alleles) such that different forms of the marker, when they are present in a homologous pair, allow transmission of each of the chromosomes in that pair to be followed. A genotype may be defined by use of one or a plurality of markers.

As used herein, "nucleic acid" includes reference to a deoxyribonucleotide or ribonucleotide polymer in either single- or double-stranded form, and unless

otherwise limited, encompasses known analogues having the essential nature of natural nucleotides in that they hybridize to single-stranded nucleic acids in a manner similar to naturally occurring nucleotides (e.g., peptide nucleic acids).

By "nucleic acid library" is meant a collection of isolated DNA or RNA molecules which comprise and substantially represent the entire transcribed fraction of a genome of a specified organism. Construction of exemplary nucleic acid libraries, such as genomic and cDNA libraries, is taught in standard molecular biology references such as Berger and Kimmel, *Guide to Molecular Cloning Techniques, Methods in Enzymology,* Vol. 152, Academic Press, Inc., San Diego, CA (Berger); Sambrook *et al., Molecular Cloning - A Laboratory Manual*, 2nd ed., Vol. 1-3 (1989); and *Current Protocols in Molecular Biology*, F.M. Ausubel *et al.*, Eds., *Current Protocols*, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc. (1994 Supplement).

As used herein "operably linked" includes reference to a functional linkage between a promoter and a second sequence, wherein the promoter sequence initiates and mediates transcription of the DNA sequence corresponding to the second sequence. Generally, operably linked means that the nucleic acid sequences being linked are contiguous and, where necessary to join two protein coding regions, contiguous and in the same reading frame.

As used herein, the term "plant" includes reference to whole plants, plant organs (e.g., leaves, stems, roots, etc.), seeds and plant cells and progeny of same. Plant cell, as used herein includes, without limitation, seeds suspension cultures, embryos, meristematic regions, callus tissue, leaves, roots, shoots, gametophytes, sporophytes, pollen, and microspores. The class of plants which can be used in the methods of the invention is generally as broad as the class of higher plants amenable to transformation techniques, including both monocotyledonous and dicotyledonous plants. A particularly preferred plant is *Zea mays*.

As used herein, "polynucleotide" includes reference to a deoxyribopolynucleotide, ribopolynucleotide, or analogs thereof, that have the essential nature of a natural ribonucleotide in that they hybridize to nucleic acids in a manner similar to naturally occurring nucleotides. A polynucleotide can be full-length or a subsequence of a native or heterologous structural or regulatory gene. Unless otherwise indicated, the term includes reference to the specified sequence

as well as the complementary sequence thereof. Thus, DNAs or RNAs with backbones modified for stability or for other reasons are "polynucleotides" as that term is intended herein. Moreover, DNAs or RNAs comprising unusual bases, such as inosine, or modified bases, such as tritylated bases, to name just two examples, are polynucleotides as the term is used herein. It will be appreciated that a great variety of modifications have been made to DNA and RNA that serve many useful purposes known to those of skill in the art. The term polynucleotide as it is employed herein embraces such chemically, enzymatically or metabolically modified forms of polynucleotides, as well as the chemical forms of DNA and RNA characteristic of viruses and cells, including *inter alia,* simple and complex cells.

The terms "polypeptide", "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical analogue of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers. The essential nature of such analogues of naturally occurring amino acids is that, when incorporated into a protein, that protein is specifically reactive to antibodies elicited to the same protein but consisting entirely of naturally occurring amino acids. The terms "polypeptide", "peptide" and "protein" are also inclusive of modifications including, but not limited to, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation. Exemplary modifications are described in most basic texts, such as, *Proteins - Structure and Molecular Properties*, 2nd ed., T. E. Creighton, W. H. Freeman and Company, New York (1993). Many detailed reviews are available on this subject, such as, for example, those provided by Wold, F., *Posttranslational Protein Modifications: Perspectives and Prospects*, pp. 1-12 in *Posttranslational Covalent Modification of Proteins*, B. C. Johnson, Ed., Academic Press, New York (1983); Seifter *et al.*, Meth. Enzymol. 182:626-646 (1990) and Rattan *et al.*, *Protein Synthesis: Posttranslational Modifications and Aging*, Ann. N.Y. Acad. Sci. 663:48-62 (1992). It will be appreciated, as is well known and as noted above, that polypeptides are not always entirely linear. For instance, polypeptides may be branched as a result of ubiquitination, and they may be circular, with or without branching, generally as a result of posttranslation events, including natural processing event and events brought about by human manipulation which do not occur naturally. Circular,

branched and branched circular polypeptides may be synthesized by non-translation natural process and by entirely synthetic methods, as well. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. In fact, blockage of the amino or carboxyl group in a polypeptide, or both, by a covalent modification, is common in naturally occurring and synthetic polypeptides and such modifications may be present in polypeptides of the present invention, as well. For instance, the amino terminal residue of polypeptides made in *E. coli* or other cells, prior to proteolytic processing, almost invariably will be N-formylmethionine. During post-translational modification of the peptide, a methionine residue at the NH_2-terminus may be deleted. Accordingly, this invention contemplates the use of both the methionine-containing and the methionineless amino terminal variants of the protein of the invention. In general, as used herein, the term polypeptide encompasses all such modifications, particularly those that are present in polypeptides synthesized by expressing a polynucleotide in a host cell.

As used herein the term "delivery protein" refers to a protein that functions in a micro-organism to deliver, integrate or otherwise move into a host cell, a region of DNA, RNA, or protein from the micro-organism. The DNA, RNA, or protein may be native to the micro-organism or foreign, or native to the host cell or foreign or foreign to both the micro-organism or host cell.

As used herein the term "pre-selected protein" refers to a deliberately chosen gene product to be expressed in a host cell.

As used herein, the term "recipient cell" refers to any host cell to be transformed in a transformation system.

As used herein, the term "co-cultivating" is used to describe the process of exposing recipient cells to a transformation vector for the purpose of introducing polynucleotides or proteins into the recipient cell. Conditions for co-cultivation will vary with the type of recipient cell and transformation vector and are known in the art.

As used herein "promoter" includes reference to a region of DNA upstream from the start of transcription and involved in recognition and binding of RNA polymerase and other proteins to initiate transcription. A "plant promoter" is a promoter capable of initiating transcription in plant cells. Exemplary plant promoters include, but are not limited to, those that are obtained from plants, plant viruses, and bacteria which comprise genes expressed in plant cells such

Agrobacterium or Rhizobium. Examples of promoters under developmental control include promoters that preferentially initiate transcription in certain tissues, such as leaves, roots, seeds, fibers, xylem vessels, tracheids, or sclerenchyma. Such promoters are referred to as "tissue preferred". Promoters which initiate transcription only in certain tissue are referred to as "tissue specific". A "cell type" specific promoter primarily drives expression in certain cell types in one or more organs, for example, vascular cells in roots or leaves. An "inducible" promoter is a promoter which is under environmental control. Examples of environmental conditions that may effect transcription by inducible promoters include anaerobic conditions or the presence of light. Tissue specific, tissue preferred, cell type specific, and inducible promoters constitute the class of "non-constitutive" promoters. A "constitutive" promoter is a promoter which is active under most environmental conditions.

The term "cell cycle polypeptide" refers to one or more amino acid sequences, in glycosylated or non-glycosylated form, involved in the regulation of cell division. The term is also inclusive of fragments, variants, homologs, alleles or precursors (e.g., preproproteins or proproteins) thereof. A "cell cycle protein" comprises a cell cycle polypeptide.

As used herein "recombinant" includes reference to a cell or vector, that has been modified by the introduction of a heterologous nucleic acid or that the cell is derived from a cell so modified. Thus, for example, recombinant cells express genes that are not found in identical form within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed or not expressed at all as a result of deliberate human intervention. The term "recombinant" as used herein does not encompass the alteration of the cell or vector by naturally occurring events (e.g., spontaneous mutation, natural transformation/transduction/transposition) such as those occurring without deliberate human intervention.

As used herein, a "recombinant expression cassette" is a nucleic acid construct, generated recombinantly or synthetically, with a series of specified nucleic acid elements which permit transcription of a particular nucleic acid in a target cell. The recombinant expression cassette can be incorporated into a plasmid, chromosome, mitochondrial DNA, plastid DNA, virus, or nucleic acid

fragment. Typically, the recombinant expression cassette portion of an expression vector includes, among other sequences, a nucleic acid to be transcribed, and a promoter.

The term "residue" or "amino acid residue" or "amino acid" are used interchangeably herein to refer to an amino acid that is incorporated into a protein, polypeptide, or peptide (collectively "protein"). The amino acid may be a naturally occurring amino acid and, unless otherwise limited, may encompass known analogs of natural amino acids that can function in a similar manner as naturally occurring amino acids.

The term "selectively hybridizes" includes reference to hybridization, under stringent hybridization conditions, of a nucleic acid sequence to a specified nucleic acid target sequence to a detectably greater degree (e.g., at least 2-fold over background) than its hybridization to non-target nucleic acid sequences and to the substantial exclusion of non-target nucleic acids. Selectively hybridizing sequences typically have about at least 80% sequence identity, preferably 90% sequence identity, and most preferably 100% sequence identity (i.e., complementary) with each other.

The term "specifically reactive", includes reference to a binding reaction between an antibody and a protein having an epitope recognized by the antigen binding site of the antibody. This binding reaction is determinative of the presence of a protein having the recognized epitope amongst the presence of a heterogeneous population of proteins and other biologics. Thus, under designated immunoassay conditions, the specified antibodies bind to an analyte having the recognized epitope to a substantially greater degree (e.g., at least 2-fold over background) than to substantially all other analytes lacking the epitope which are present in the sample.

Specific binding to an antibody under such conditions may require an antibody that is selected for its specificity for a particular protein. For example, antibodies raised to the polypeptides of the present invention can be selected from those antibodies that are specifically reactive with polypeptides of the present invention. The proteins used as immunogens can be in native conformation or denatured so as to provide a linear epitope.

A variety of immunoassay formats may be used to select antibodies specifically reactive with a particular protein (or other analyte). For example, solid-

phase ELISA immunoassays are routinely used to select monoclonal antibodies specifically immunoreactive with a protein. See Harlow and Lane, *Antibodies, A Laboratory Manual*, Cold Spring Harbor Publications, New York (1988), for a description of immunoassay formats and conditions that can be used to determine selective reactivity.

The terms "stringent conditions" or "stringent hybridization conditions" includes reference to conditions under which a probe will hybridize to its target sequence, to a detectably greater degree than other sequences (e.g., at least 2-fold over background). Stringent conditions are sequence-dependent and will be different in different circumstances. By controlling the stringency of the hybridization and/or washing conditions, target sequences can be identified which are 100% complementary to the probe (homologous probing). Alternatively, stringency conditions can be adjusted to allow some mismatching in sequences so that lower degrees of similarity are detected (heterologous probing). Generally, a probe is less than about 1000 nucleotides in length, preferably less than 500 nucleotides in length.

Typically, stringent conditions will be those in which the salt concentration is less than about 1.5 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (e.g., 10 to 50 nucleotides) and at least about 60°C for long probes (e.g., greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. Exemplary low stringency conditions include hybridization with a buffer solution of 30 to 35% formamide, 1 M NaCl, 1% SDS (sodium dodecyl sulfate) at 37°C, and a wash in 1X to 2X SSC (20X SSC = 3.0 M NaCl/0.3 M trisodium citrate) at 50 to 55°C. Exemplary moderate stringency conditions include hybridization in 40 to 45% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.5X to 1X SSC at 55 to 60°C. Exemplary high stringency conditions include hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.1X SSC at 60 to 65°C.

Specificity is typically the function of post-hybridization washes, the critical factors being the ionic strength and temperature of the final wash solution. For DNA-DNA hybrids, the T_m can be approximated from the equation of Meinkoth and Wahl, Anal. Biochem., 138:267-284 (1984): T_m = 81.5 °C + 16.6 (log M) + 0.41 (%GC) - 0.61 (% form) - 500/L; where M is the molarity of monovalent cations,

%GC is the percentage of guanosine and cytosine nucleotides in the DNA, % form is the percentage of formamide in the hybridization solution, and L is the length of the hybrid in base pairs. The T_m is the temperature (under defined ionic strength and pH) at which 50% of a complementary target sequence hybridizes to a perfectly matched probe. T_m is reduced by about 1 °C for each 1% of mismatching; thus, T_m, hybridization and/or wash conditions can be adjusted to hybridize to sequences of the desired identity. For example, if sequences with $\geq$90% identity are sought, the T_m can be decreased 10 °C. Generally, stringent conditions are selected to be about 5 °C lower than the thermal melting point (T_m) for the specific sequence and its complement at a defined ionic strength and pH. However, severely stringent conditions can utilize a hybridization and/or wash at 1, 2, 3, or 4 °C lower than the thermal melting point (T_m); moderately stringent conditions can utilize a hybridization and/or wash at 6, 7, 8, 9, or 10 °C lower than the thermal melting point (T_m); low stringency conditions can utilize a hybridization and/or wash at 11, 12, 13, 14, 15, or 20 °C lower than the thermal melting point (T_m). Using the equation, hybridization and wash compositions, and desired T_m, those of ordinary skill will understand that variations in the stringency of hybridization and/or wash solutions are inherently described. If the desired degree of mismatching results in a T_m of less than 45 °C (aqueous solution) or 32 °C (formamide solution) it is preferred to increase the SSC concentration so that a higher temperature can be used. An extensive guide to the hybridization of nucleic acids is found in Tijssen, *Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes*, Part I, Chapter 2 "*Overview of principles of hybridization and the strategy of nucleic acid probe assays*", Elsevier, New York (1993); and *Current Protocols in Molecular Biology*, Chapter 2, Ausubel, *et al.*, Eds., Greene Publishing and Wiley-Interscience, New York (1995).

As used herein, "transgenic plant" includes reference to a plant which comprises within its genome a heterologous polynucleotide. Generally, the heterologous polynucleotide is stably integrated within the genome such that the polynucleotide is passed on to successive generations. The heterologous polynucleotide may be integrated into the genome alone or as part of a recombinant expression cassette. "Transgenic" is used herein to include any cell, cell line, callus, tissue, plant part or plant, the genotype of which has been altered

by the presence of heterologous nucleic acid including those transgenics initially so altered as well as those created by sexual crosses or asexual propagation from the initial transgenic. The term "transgenic" as used herein does not encompass the alteration of the genome (chromosomal or extra-chromosomal) by conventional plant breeding methods or by naturally occurring events such as random cross-fertilization, non-recombinant viral infection, non-recombinant bacterial transformation, non-recombinant transposition, or spontaneous mutation.

As used herein, "vector" includes reference to a nucleic acid used in transfection of a host cell and into which can be inserted a polynucleotide. Vectors are often replicons. Expression vectors permit transcription of a nucleic acid inserted therein.

The following terms are used to describe the sequence relationships between two or more nucleic acids or polynucleotides: (a) "reference sequence", (b) "comparison window", (c) "sequence identity", (d) "percentage of sequence identity", and (e) "substantial identity".

(a) As used herein, "reference sequence" is a defined sequence used as a basis for sequence comparison. A reference sequence may be a subset or the entirety of a specified sequence; for example, as a segment of a full-length cDNA or gene sequence, or the complete cDNA or gene sequence.

(b) As used herein, "comparison window" means includes reference to a contiguous and specified segment of a polynucleotide sequence, wherein the polynucleotide sequence may be compared to a reference sequence and wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Generally, the comparison window is at least 20 contiguous nucleotides in length, and optionally can be 30, 40, 50, 100, or longer. Those of skill in the art understand that to avoid a high similarity to a reference sequence due to inclusion of gaps in the polynucleotide sequence a gap penalty is typically introduced and is subtracted from the number of matches.

Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison may be conducted by the local homology algorithm of Smith and Waterman, Adv. Appl. Math. 2:482 (1981); by the homology alignment algorithm of Needleman and Wunsch, J. Mol. Biol.

48:443 (1970); by the search for similarity method of Pearson and Lipman, Proc. Natl. Acad. Sci. 85:2444 (1988); by computerized implementations of these algorithms, including, but not limited to: CLUSTAL in the PC/Gene program by Intelligenetics, Mountain View, California, GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, Wisconsin, USA; the CLUSTAL program is well described by Higgins and Sharp, Gene 73:237-244 (1988); Higgins and Sharp, CABIOS 5:151-153 (1989); Corpet, *et al.*, Nucleic Acids Research 16:10881-90 (1988); Huang, *et al.*, Computer Applications in the Biosciences 8:155-65 (1992), and Pearson, *et al.*, Methods in Molecular Biology 24:307-331 (1994). The BLAST family of programs which can be used for database similarity searches includes: BLASTN for nucleotide query sequences against nucleotide database sequences; BLASTX for nucleotide query sequences against protein database sequences; BLASTP for protein query sequences against protein database sequences; TBLASTN for protein query sequences against nucleotide database sequences; and TBLASTX for nucleotide query sequences against nucleotide database sequences. See, *Current Protocols in Molecular Biology*, Chapter 19, Ausubel, *et al.*, Eds., Greene Publishing and Wiley-Interscience, New York (1995).

For purposes of defining the present invention, the Gap 10 program in the Wisconsin Genetics Software Package using default parameters is used, Genetics Computer Group (GCG), 575 Science Dr., Madison, Wisconsin, USA. The algorithm used for the GAP program is that of Needleman and Wunsch (J. Mol. Biol. 48:443-453 [1970]). The parameters used are as follows: for nucleotide comparisons the gap creation penalty = 50, gap extension penalty = 3; for amino acid comparisons the gap creation penalty = 12, the gap extension penalty = 4.

As those of ordinary skill in the art will understand, searches assume that proteins can be modeled as random sequences. However, many real proteins comprise regions of nonrandom sequences which may be homopolymeric tracts, short-period repeats, or regions enriched in one or more amino acids. Such low-complexity regions may be aligned between unrelated proteins even though other regions of the protein are entirely dissimilar. A number of low-complexity filter programs can be employed to reduce such low-complexity alignments. For example, the SEG (Wooten and Federhen, *Comput. Chem.*, 17:149-163 (1993))

and XNU (Claverie and States, *Comput. Chem.*, 17:191-201 (1993)) low-complexity filters can be employed alone or in combination.

(c) As used herein, "sequence identity" or "identity" in the context of two nucleic acid or polypeptide sequences includes reference to the residues in the two sequences which are the same when aligned for maximum correspondence over a specified comparison window. When percentage of sequence identity is used in reference to proteins it is recognized that residue positions which are not identical often differ by conservative amino acid substitutions, where amino acid residues are substituted for other amino acid residues with similar chemical properties (e.g. charge or hydrophobicity) and therefore do not change the functional properties of the molecule. Where sequences differ in conservative substitutions, the percent sequence identity may be adjusted upwards to correct for the conservative nature of the substitution. Sequences which differ by such conservative substitutions are said to have "sequence similarity" or "similarity". Means for making this adjustment are well-known to those of skill in the art. Typically this involves scoring a conservative substitution as a partial rather than a full mismatch, thereby increasing the percentage sequence identity. Thus, for example, where an identical amino acid is given a score of 1 and a non-conservative substitution is given a score of zero, a conservative substitution is given a score between zero and 1. The scoring of conservative substitutions is calculated, *e.g.*, according to the algorithm of Meyers and Miller, Computer Applic. Biol. Sci., 4:11-17 (1988) e.g., as implemented in the program PC/GENE (Intelligenetics, Mountain View, California, USA).

(d) As used herein, "percentage of sequence identity" means the value determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity. For purposes of defining the invention, % identity on the nucleic acid level

is determined by the BESTFIT DNA Sequence Alignment software on Genescape using a gap weight of 50 and a length weight of 3. For purposes of defining the invention, % identity on the amino acid level is determined by the BESTFIT DNA Sequence Alignment software on Genescape using a gap weight of 12 and a length weight of 4.

(e) (i) The term "substantial identity" of polynucleotide sequences means that a polynucleotide comprises a sequence that has at least 70% sequence identity, preferably at least 80%, more preferably at least 90% and most preferably at least 95%, compared to a reference sequence using one of the alignment programs described using standard parameters. One of skill will recognize that these values can be appropriately adjusted to determine corresponding identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning and the like. Substantial identity of amino acid sequences for these purposes normally means sequence identity of at least 85%, more preferably at least 90%, and most preferably at least 98%.

Another indication that nucleotide sequences are substantially identical is if two molecules hybridize to each other under stringent conditions. However, nucleic acids which do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This may occur, e.g., when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. One indication that two nucleic acid sequences are substantially identical is that the polypeptide which the first nucleic acid encodes is immunologically cross reactive with the polypeptide encoded by the second nucleic acid.

(e) (ii) The terms "substantial identity" in the context of a peptide indicates that a peptide comprises a sequence with at least 70% sequence identity to a reference sequence, preferably 80%, more preferably 85%, most preferably at least 90% or 95% sequence identity to the reference sequence over a specified comparison window. Optimal alignment can be conducted using the homology alignment algorithm of Needleman and Wunsch, J. Mol. Biol. 48:443 (1970). An indication that two peptide sequences are substantially identical is that one peptide is immunologically reactive with antibodies raised against the second peptide. Thus, a peptide is substantially identical to a second peptide, for

example, where the two peptides differ only by a conservative substitution. Peptides which are "substantially similar" share sequences as noted above except that residue positions which are not identical may differ by conservative amino acid changes.

By "two-hybrid system" is meant a screening method to identify protein-protein interactions, using a known gene (and its encoded product) as a "bait" or target and screening a library of expressed genes and their corresponding encoded products for specific interactions with the "bait" molecule. Methods for library construction and use of visual marker genes for yeast two-hybrid screens are well known in the art, and can be found in Sambrook, *et al.*, 1990, Ausubel *et al.*, 1990 and G. Hannon and P. Bartel, Identification of interacting proteins using the two-hybrid system. Methods Mol. Cellular Biol. 5:289-297 (1995).

By "functionally equivalent" is intended that the sequence of the variant defines a chain that produces a protein having substantially the same biological effect as the native protein of interest. The variant is catalytically active.

By "modulate" is intended to increase, decrease, influence or change.

By "Cataytically Active" is intended the ability of a protein to bind to retinoblastoma (Rb) or is involved in stimulating DNA replication during the cell cycle.

DETAILED DESCRIPTION OF THE INVENTION

DP is a transcription factor that activates transcription of numerous genes involved in DNA replication, thus playing a role in the G1 to S transition in the cell cycle. DP, also called DRTF (differentiation-regulated transcription factor 1), has been shown to form specific multiprotein complexes with the retinoblastoma susceptibility protein (Rb), p107 (a homologue of Rb), cyclins and cdk2 (Bagehi *et al.*, 1991; Pandara and La Thangue, 1991; Chelloppan *et al.*, 1991; Chittenden *et al.*, 1991). DP was first identified as a cellular transcription factor that could bind and activate the adenovirus EZ promoter.

The present invention provides, *inter alia,* compositions and methods for modulating (i.e., increasing or decreasing) the total levels of proteins of the present invention and/or altering their ratios in plants. Thus, the present invention provides utility in such exemplary applications as the regulation of cell division.

The polypeptides of the present invention can be expressed at times or in quantities which are not characteristic of non-recombinant plants.

In particular, modulating cell cycle proteins is expected to provide a positive growth advantage and increase crop yield. Cell cycle nucleic acids can be adducted to a second nucleic acid sequence encoding a DNA-binding domain, for use in two-hybrid systems to identify DP-interacting proteins. It is expected that modulating the level of cell cycle protein, i.e. overexpression, will increase endoreduplication which is expected to increase the size of the seed, the size of the endosperm and amount of protein in the seed. The cell cycle protein can be used to affinity purify Rb or associated components.

The present invention also provides isolated nucleic acid comprising polynucleotides of sufficient length and complementarity to a cell cycle gene to use as probes or amplification primers in the detection, quantitation, or isolation of gene transcripts. For example, isolated nucleic acids of the present invention can be used as probes in detecting deficiencies in the level of mRNA in screenings for desired transgenic plants, for detecting mutations in the gene (e.g., substitutions, deletions, or additions), for monitoring upregulation of expression or changes in enzyme activity in screening assays of compounds, for detection of any number of allelic variants (polymorphisms) of the gene, or for use as molecular markers in plant breeding programs. The isolated nucleic acids of the present invention can also be used for recombinant expression of cell cycle polypeptides, or for use as immunogens in the preparation and/or screening of antibodies. The isolated nucleic acids of the present invention can also be employed for use in sense or antisense suppression of one or more cell cycle genes in a host cell, tissue, or plant. Attachment of chemical agents which bind, intercalate, cleave and/or crosslink to the isolated nucleic acids of the present invention can also be used to modulate transcription or translation. Further, using a primer specific to an insertion sequence (e.g., transposon) and a primer which specifically hybridizes to an isolated nucleic acid of the present invention, one can use nucleic acid amplification to identity insertion sequence inactivated cell cycle genes from a cDNA library prepared from insertion sequence mutagenized plants. Progeny seed from the plants comprising the desired inactivated gene can be grown to a plant to study the phenotypic changes characteristic of that inactivation. See, Tools to Determine the Function of Genes, 1995 Proceedings of the Fiftieth

Annual Corn and Sorghum Industry Research Conference, American Seed Trade Association, Washington, D.C., 1995. Additionally, non-translated 5' or 3' regions of the polynucleotides of the present invention can be used to modulate turnover of heterologous mRNAs and/or protein synthesis. Further, the codon preference characteristic of the polynucleotides of the present invention can be employed in heterologous sequences, or altered in homologous or heterologous sequences, to modulate translational level and/or rates.

The present invention also provides isolated proteins comprising polypeptides including an amino acid sequence from the cell cycle polypeptides (e.g., preproenzyme, proenzyme, or enzymes) as disclosed herein. The present invention also provides proteins comprising at least one epitope from a cell cycle polypeptide. The proteins of the present invention can be employed in assays for enzyme agonists or antagonists of enzyme function, or for use as immunogens or antigens to obtain antibodies specifically immunoreactive with a protein of the present invention. Such antibodies can be used in assays for expression levels, for identifying and/or isolating nucleic acids of the present invention from expression libraries, or for purification of cell cycle polypeptides.

The isolated nucleic acids of the present invention can be used over a broad range of plant types, including species from the genera *Cucurbita, Rosa, Vitis, Juglans, Fragaria, Lotus, Medicago, Onobrychis, Trifolium, Trigonella, Vigna, Citrus, Linum, Geranium, Manihot, Daucus, Arabidopsis, Brassica, Raphanus, Sinapis, Atropa, Capsicum, Datura, Hyoscyamus, Lycopersicon, Nicotiana, Solanum, Petunia, Digitalis, Majorana, Ciahorium, Helianthus, Lactuca, Bromus, Asparagus, Antirrhinum, Heterocallis, Nemesis, Pelargonium, Panieum, Pennisetum, Ranunculus, Senecio, Salpiglossis, Cucumis, Browaalia, Glycine, Pisum, Phaseolus, Lolium, Oryza, Zea, Avena, Hordeum, Secale, Triticum, Sorghum, Picea, and Populus*. Preferred plants include corn, soybeans, sorghum, sunflower, wheat, rice, alfalfa and canola.

NUCLEIC ACIDS

The isolated nucleic acids of the present invention can be made using (a) standard recombinant methods, (b) synthetic techniques, or combinations thereof. In some embodiments, the polynucleotides of the present invention will be cloned, amplified, or otherwise constructed from a monocot or dicot. In preferred

embodiments the monocot is corn, sorghum, barley, wheat, millet, or rice. Preferred dicots include soybeans, sunflower, canola, alfalfa, cotton, potato, or cassava.

Functional fragments included in the invention can be obtained using primers that selectively hybridize under stringent conditions. Fragments can be made through site directed mutagenesis, restriction, change, DNA shuffling or a variety of methods known in the art. Primers are generally at least 12 bases in length and can be as high as 200 bases, but will generally be from 15 to 75, preferably from 15 to 50. Functional fragments can be identified using a variety of techniques such as restriction analysis, Southern analysis, primer extension analysis, and DNA sequence analysis and then tested for catalytic activity.

The present invention includes a plurality of polynucleotides that encode for the identical amino acid sequence. The degeneracy of the genetic code allows for such "silent variations" which can be used, for example, to selectively hybridize and detect allelic variants of polynucleotides of the present invention. Additionally, the present invention includes isolated nucleic acids comprising allelic variants. The term "allele" as used herein refers to a related nucleic acid of the same gene.

Variants of nucleic acids included in the invention can be obtained, for example, by oligonucleotide-directed mutagenesis, linker-scanning mutagenesis, mutagenesis using the polymerase chain reaction, and the like. See, for example, Ausubel, pages 8.0.3 - 8.5.9. Also, see generally, McPherson (ed.), *DIRECTED MUTAGENESIS: A Practical approach*, (IRL Press, 1991). Thus, the present invention also encompasses DNA molecules comprising nucleotide sequences that have substantial sequence similarity with the inventive sequences.

Variants included in the invention may contain individual substitutions, deletions or additions to the nucleic acid or polypeptide sequences. Such changes will alter, add or delete a single amino acid or a small percentage of amino acids in the encoded sequence. Variants are referred to as "conservatively modified variants" where the alteration results in the substitution of an amino acid with a chemically similar amino acid. When the nucleic acid is prepared or altered synthetically, advantage can be taken of known codon preferences of the intended host.

The present invention also includes "shufflents" produced by sequence shuffling of the inventive polynucleotides to obtain a desired characteristic.

Sequence shuffling is described in PCT publication No. 96/19256. See also, Zhang, J.- H., *et al. Proc. Natl. Acad. Sci. USA* 94:4504-4509 (1997).

The present invention also includes the use of 5' and/or 3' UTR regions for modulation of translation of heterologous coding sequences. Positive sequence motifs include translational initiation consensus sequences (Kozak, *Nucleic Acids Res.*15:8125 (1987)) and the 7-methylguanosine cap structure (Drummond *et al., Nucleic Acids Res.* 13:7375 (1985)). Negative elements include stable intramolecular 5' UTR stem-loop structures (Muesing *et al., Cell* 48:691 (1987)) and AUG sequences or short open reading frames preceded by an appropriate AUG in the 5' UTR (Kozak, *supra, Rao et al., Mol. and Cell. Biol.* 8:284 (1988)).

Further, the polypeptide-encoding segments of the polynucleotides of the present invention can be modified to alter codon usage. Altered codon usage can be employed to alter translational efficiency and/or to optimize the coding sequence for expression in a desired host or to optimize the codon usage in a heterologous sequence for expression in maize. Codon usage in the coding regions of the polynucleotides of the present invention can be analyzed statistically using commercially available software packages such as "Codon Preference" available from the University of Wisconsin Genetics Computer Group (see Devereaux *et al., Nucleic Acids Res.* 12: 387-395 (1984)) or MacVector 4.1 (Eastman Kodak Co., New Haven, Conn.).

For example, the inventive nucleic acids can be optimized for enhanced or suppressed expression in organisms of interest. See, for example, EPA0359472; WO91/16432; Perlak *et al.* (1991) *Proc. Natl. Acad. Sci. USA 88*:3324-3328; and Murray *et al.* (1989) *Nucleic Acids Res. 17*:477-498. In this manner, the genes can be synthesized utilizing species-preferred codons. See, for example, Murray *et al.* (1989) *Nucleic Acids Res. 17*:477-498, the disclosure of which is incorporated herein by reference.

The present invention provides subsequences comprising isolated nucleic acids containing at least 16 contiguous bases of the inventive sequences. For example the isolated nucleic acid includes those comprising at least 20, 25, 30, 40, 50, 60, 75 or 100 contiguous nucleotides of the inventive sequences. Subsequences of the isolated nucleic acid can be used to modulate or detect gene expression by introducing into the subsequences compounds which bind, intercalate, cleave and/or crosslink to nucleic acids.

The nucleic acids of the invention may conveniently comprise a multi-cloning site comprising one or more endonuclease restriction sites inserted into the nucleic acid to aid in isolation of the polynucleotide. Also, translatable sequences may be inserted to aid in the isolation of the translated polynucleotide of the present invention. For example, a hexa-histidine marker sequence provides a convenient means to purify the proteins of the present invention.

A polynucleotide of the present invention can be attached to a vector, adapter, promoter, transit peptide or linker for cloning and/or expression of a polynucleotide of the present invention. Additional sequences may be added to such cloning and/or expression sequences to optimize their function in cloning and/or expression, to aid in isolation of the polynucleotide, or to improve the introduction of the polynucleotide into a cell. Use of cloning vectors, expression vectors, adapters, and linkers is well known and extensively described in the art. For a description of such nucleic acids see, for example, Stratagene Cloning Systems, Catalogs 1995, 1996, 1997 (La Jolla, CA); and, Amersham Life Sciences, Inc, Catalog '97 (Arlington Heights, IL).

The isolated nucleic acid compositions of this invention, such as RNA, cDNA, genomic DNA, or a hybrid thereof, can be obtained from plant biological sources using any number of cloning methodologies known to those of skill in the art. In some embodiments, oligonucleotide probes which selectively hybridize, under stringent conditions, to the polynucleotides of the present invention are used to identify the desired sequence in a cDNA or genomic DNA library.

Exemplary total RNA and mRNA isolation protocols are described in *Plant Molecular Biology: A Laboratory Manual*, Clark, Ed., Springer-Verlag, Berlin (1997); and, *Current Protocols in Molecular Biology*, Ausubel, *et al.*, Eds., Greene Publishing and Wiley-Interscience, New York (1995). Total RNA and mRNA isolation kits are commercially available from vendors such as Stratagene (La Jolla, CA), Clonetech (Palo Alto, CA), Pharmacia (Piscataway, NJ), and 5'-3' (Paoli, PA). See also, U.S. Patent Nos. 5,614,391; and, 5,459,253.

Typical cDNA synthesis protocols are well known to the skilled artisan and are described in such standard references as: *Plant Molecular Biology: A Laboratory Manual*, Clark, Ed., Springer-Verlag, Berlin (1997); and, *Current Protocols in Molecular Biology*, Ausubel, *et al.*, Eds., Greene Publishing and

Wiley-Interscience, New York (1995). cDNA synthesis kits are available from a variety of commercial vendors such as Stratagene or Pharmacia.

An exemplary method of constructing a greater than 95% pure full-length cDNA library is described by Carninci *et al., Genomics*, 37:327-336 (1996). Other methods for producing full-length libraries are known in the art. See, e.g., Edery *et al., Mol. Cell Biol.*,15(6):3363-3371 (1995); and, PCT Application WO 96/34981.

It is often convenient to normalize a cDNA library to create a library in which each clone is more equally represented. A number of approaches to normalize cDNA libraries are known in the art. Construction of normalized libraries is described in Ko, *Nucl. Acids. Res.*, 18(19):5705-5711 (1990); Patanjali *et al., Proc. Natl. Acad. U.S.A.*, 88:1943-1947 (1991); U.S. Patents 5,482,685 and 5,637,685; and Soares *et al.*, *Proc. Natl. Acad. Sci. USA*, 91:9228-9232 (1994).

Subtracted cDNA libraries are another means to increase the proportion of less abundant cDNA species. See, *Foote et al.* in, *Plant Molecular Biology: A Laboratory Manual*, Clark, Ed., Springer-Verlag, Berlin (1997); Kho and Zarbl, *Technique*, 3(2):58-63 (1991); Sive and St. John, *Nucl. Acids Res.*, 16(22):10937 (1988); *Current Protocols in Molecular Biology*, Ausubel, *et al.*, Eds., Greene Publishing and Wiley-Interscience, New York (1995); and, Swaroop *et al., Nucl. Acids Res.*, 19)8):1954 (1991). cDNA subtraction kits are commercially available. See, e.g., PCR-Select (Clontech).

To construct genomic libraries, large segments of genomic DNA are generated by random fragmentation. Examples of appropriate molecular biological techniques and instructions are found in Sambrook, *et al.*, *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Vols. 1-3 (1989), Methods in Enzymology, Vol. 152: *Guide to Molecular Cloning Techniques*, Berger and Kimmel, Eds., San Diego: Academic Press, Inc. (1987), *Current Protocols in Molecular Biology*, Ausubel, *et al.*, Eds., Greene Publishing and Wiley-Interscience, New York (1995); *Plant Molecular Biology: A Laboratory Manual*, Clark, Ed., Springer-Verlag, Berlin (1997). Kits for construction of genomic libraries are also commercially available.

The cDNA or genomic library can be screened using a probe based upon the sequence of a nucleic acid of the present invention such as those disclosed herein. Probes may be used to hybridize with genomic DNA or cDNA sequences to isolate homologous genes in the same or different plant species. Those of skill

in the art will appreciate that various degrees of stringency of hybridization can be employed in the assay; and either the hybridization or the wash medium can be stringent. The degree of stringency can be controlled by temperature, ionic strength, pH and the presence of a partially denaturing solvent such as formamide.

Typically, stringent hybridization conditions will be those in which the salt concentration is less than about 1.5 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (e.g., 10 to 50 nucleotides) and at least about 60°C for long probes (e.g., greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide.
Preferably the hybridization is conducted under low stringency conditions which include hybridization with a buffer solution of 30 % formamide, 1 M NaCl, 1% SDS (sodium dodecyl sulfate) at 37°C, and a wash in 1X to 2X SSC (20X SSC = 3.0 M NaCl/0.3 M trisodium citrate) at 50°C. More preferably the hybridization is conducted under moderate stringency conditions which include hybridization in 40 % formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.5X to 1X SSC at 55°C. Most preferably the hybridization is conducted under high stringency conditions which include hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.1X SSC at 60°C.

The time for conducting the hybridization is not critical and can be in the range of from 4 to 16 hours.

An extensive guide to the hybridization of nucleic acids is found in Tijssen, *Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes,* Part I, Chapter 2 "*Overview of principles of hybridization and the strategy of nucleic acid probe assays*", Elsevier, New York (1993); and *Current Protocols in Molecular Biology*, Chapter 2, Ausubel, *et al.*, Eds., Greene Publishing and Wiley-Interscience, New York (1995). Often, cDNA libraries will be normalized to increase the representation of relatively rare cDNAs.

The nucleic acids of the invention can be amplified from nucleic acid samples using amplification techniques. For instance, polymerase chain reaction (PCR) technology can be used to amplify the sequences of polynucleotides of the present invention and related genes directly from genomic DNA or cDNA libraries. PCR and other *in vitro* amplification methods may also be useful, for example, to

clone nucleic acid sequences that code for proteins to be expressed, to make nucleic acids to use as probes for detecting the presence of the desired mRNA in samples, for nucleic acid sequencing, or for other purposes.

Examples of techniques useful for *in vitro* amplification methods are found in Berger, Sambrook, and Ausubel, as well as Mullis *et al.*, U.S. Patent No. 4,683,202 (1987); and, *PCR Protocols A Guide to Methods and Applications*, Innis *et al.*, Eds., Academic Press Inc., San Diego, CA (1990). Commercially available kits for genomic PCR amplification are known in the art. See, e.g., Advantage-GC Genomic PCR Kit (Clontech). The T4 gene 32 protein (Boehringer Mannheim) can be used to improve yield of long PCR products.

PCR-based screening methods have also been described. Wilfinger *et al.* describe a PCR-based method in which the longest cDNA is identified in the first step so that incomplete clones can be eliminated from study. *BioTechniques*, 22(3): 481-486 (1997).

In one aspect of the invention, nucleic acids can be amplified from a *Zea mays* nucleic acid library. The nucleic acid library may be a cDNA library, a genomic library, or a library generally constructed from nuclear transcripts at any stage of intron processing.

Libraries can be made from a variety of maize tissues. Good results have been obtained using mitotically active tissues such as shoot meristems, shoot meristem cultures, embryos, callus and suspension cultures, immature ears and tassels, and young seedlings. The cDNA of the present invention was obtained from developing endosperm. Since cell cycle proteins are typically expressed at specific cell cycle stages it may be possible to enrich for such rare messages using exemplary cell cycle inhibitors such as aphidicolin, hydroxyurea, mimosine, and double-phosphate starvation methods to block cells at the G1/S boundary. Cells can also be blocked at this stage using the double phosphate starvation method. Hormone treatments that stimulate cell division, for example cytokinin, would also increase expression of the cell cycle RNA.

Alternatively, the sequences of the invention can be used to isolate corresponding sequences in other organisms, particularly other plants, more particularly, other monocots. In this manner, methods such as PCR, hybridization, and the like can be used to identify such sequences having substantial sequence similarity to the sequences of the invention. See, for example, Sambrook *et al.*

(1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Plainview, New York). and Innis *et al.* (1990), *PCR Protocols: A Guide to Methods and Applications* (Academic Press, New York). Coding sequences isolated based on their sequence identity to the entire inventive coding sequences set forth herein or to fragments thereof are encompassed by the present invention.

The isolated nucleic acids of the present invention can also be prepared by direct chemical synthesis by methods such as the phosphotriester method of Narang *et al., Meth. Enzymol.* 68: 90-99 (1979); the phosphodiester method of Brown *et al., Meth. Enzymol.* 68: 109-151 (1979); the diethylphosphoramidite method of Beaucage *et al., Tetra. Lett.* 22: 1859-1862 (1981); the solid phase phosphoramidite triester method described by Beaucage and Caruthers, *Tetra. Letts.* 22(20): 1859-1862 (1981), *e.g.*, using an automated synthesizer, *e.g.*, as described in Needham-VanDevanter *et al., Nucleic Acids Res.*, 12: 6159-6168 (1984); and, the solid support method of U.S. Patent No. 4,458,066. Chemical synthesis generally produces a single stranded oligonucleotide. This may be converted into double stranded DNA by hybridization with a complementary sequence, or by polymerization with a DNA polymerase using the single strand as a template. One of skill will recognize that while chemical synthesis of DNA is limited to sequences of about 100 bases, longer sequences may be obtained by the ligation of shorter sequences.

EXPRESSION CASSETTES

In another embodiment expression cassettes comprising isolated nucleic acids of the present invention are provided. An expression cassette will typically comprise a polynucleotide of the present invention operably linked to transcriptional initiation regulatory sequences which will direct the transcription of the polynucleotide in the intended host cell, such as tissues of a transformed plant, bacterial or yeast hosts.

The construction of expression cassettes that can be employed in conjunction with the present invention is well known to those of skill in the art in light of the present disclosure. See, e.g., Sambrook, *et al.*; Molecular Cloning: A Laboratory Manual; Cold Spring Harbor, New York; (1989); Gelvin, *et al.*; Plant Molecular Biology Manual; (1990); Plant Biotechnology: Commercial Prospects

and Problems, eds. Prakash, *et al.*; Oxford & IBH Publishing Co.; New Delhi, India; (1993); and Heslot, *et al.*; Molecular Biology and Genetic Engineering of Yeasts; CRC Press, Inc., USA; (1992); each incorporated herein in its entirety by reference. For example, plant expression vectors may include (1) a cloned plant nucleic acid under the transcriptional control of 5' and 3' regulatory sequences and (2) a dominant selectable marker. Such plant expression vectors may also contain, if desired, a promoter regulatory region (e.g., one conferring inducible, constitutive, environmentally- or developmentally-regulated, or cell- or tissue-specific/selective expression), a transcription initiation start site, a ribosome binding site, an RNA processing signal, a transcription termination site, and/or a polyadenylation signal.

Constitutive, tissue-preferred or inducible promoters can be employed. Examples of constitutive promoters include the cauliflower mosaic virus (CaMV) 35S transcription initiation region, the 1'- or 2'- promoter derived from T-DNA of *Agrobacterium tumefaciens*, the ubiquitin 1 promoter, the Smas promoter, the cinnamyl alcohol dehydrogenase promoter (U.S. Patent No. 5,683,439), the *Nos* promoter, the pEmu promoter, the rubisco promoter, the GRP1-8 promoter and other transcription initiation regions from various plant genes known to those of skill.

Examples of inducible promoters are the Adh1 promoter which is inducible by hypoxia or cold stress, the Hsp70 promoter which is inducible by heat stress, and the PPDK promoter which is inducible by light. Also useful are promoters which are chemically inducible.

Examples of promoters under developmental control include promoters that initiate transcription preferentially in certain tissues, such as leaves, roots, fruit, seeds, or flowers. An exemplary promoter is the anther specific promoter 5126 (U.S. Patent Nos. 5,689,049 and 5,689,051). Examples of seed-preferred promoters include, but are not limited to, 27 kD gamma zein promoter and waxy promote, Boronat,A., Martinez,M.C., Reina,M., Puigdomenech,P. and Palau,J.; Isolation and sequencing of a 28 kD glutelin-2 gene from maize: Common elements in the 5' flanking regions among zein and glutelin genes; Plant Sci. 47, 95-102 (1986), and Reina,M., Ponte,I., Guillen,P., Boronat,A. and Palau,J., Sequence analysis of a genomic clone encoding a Zc2 protein from Zea mays W64 A, Nucleic Acids Res. 18 (21), 6426 (1990). See the following site relating to

the waxy promoter: Kloesgen,R.B., Gierl,A., Schwarz-Sommer,ZS. and Saedler,H., Molecular analysis of the waxy locus of Zea mays, Mol. Gen. Genet. 203, 237-244 (1986). Promoters that express in the embryo, pericarp, and endosperm are also known.

Either heterologous or non-heterologous (i.e., endogenous) promoters can be employed to direct expression of the nucleic acids of the present invention. These promoters can also be used, for example, in expression cassettes to drive expression of antisense nucleic acids to reduce, increase, or alter concentration and/or composition of the proteins of the present invention in a desired tissue.

If polypeptide expression is desired, it is generally desirable to include a polyadenylation region at the 3'-end of a polynucleotide coding region. The polyadenylation region can be derived from the natural gene, from a variety of other plant genes, or from T-DNA. The 3' end sequence to be added can be derived from, for example, the nopaline synthase or octopine synthase genes, or alternatively from another plant gene, or less preferably from any other eukaryotic gene.

An intron sequence can be added to the 5' untranslated region or the coding sequence of the partial coding sequence to increase the amount of the mature message that accumulates. See for example Buchman and Berg, *Mol. Cell Biol.* 8: 4395-4405 (1988); Callis *et al.*, *Genes Dev.* 1: 1183-1200 (1987). Use of maize introns Adh1-S intron 1, 2, and 6, the Bronze-1 intron are known in the art. See generally, *The Maize Handbook*, Chapter 116, Freeling and Walbot, Eds., Springer, New York (1994).

The vector comprising the sequences from a polynucleotide of the present invention will typically comprise a marker gene which confers a selectable phenotype on plant cells. Usually, the selectable marker gene will encode antibiotic or herbicide resistance. Suitable genes include those coding for resistance to the antibiotic spectinomycin or streptomycin (e.g., the aada gene), the streptomycin phosphotransferase (SPT) gene coding for streptomycin resistance, the neomycin phosphotransferase (NPTII) gene encoding kanamycin or geneticin resistance, the hygromycin phosphotransferase (HPT) gene coding for hygromycin resistance.

Suitable genes coding for resistance to herbicides include those which act to inhibit the action of acetolactate synthase (ALS), in particular the

sulfonylurea-type herbicides (e.g., the acetolactate synthase (ALS) gene containing mutations leading to such resistance in particular the S4 and/or Hra mutations), those which act to inhibit action of glutamine synthase, such as phosphinothricin or basta (e.g., the *bar* gene), or other such genes known in the art. The *bar* gene encodes resistance to the herbicide basta and the ALS gene encodes resistance to the herbicide chlorsulfuron.

Typical vectors useful for expression of nucleic acids in higher plants are well known in the art and include vectors derived from the tumor-inducing (Ti) plasmid of *Agrobacterium tumefaciens* described by Rogers *et al.*, Meth. In Enzymol., 153:253-277 (1987). Exemplary *A. tumefaciens* vectors useful herein are plasmids pKYLX6 and pKYLX7 of Schardl *et al.*, Gene, 61:1-11 (1987) and Berger *et al.*, Proc. Natl. Acad. Sci. U.S.A., 86:8402-8406 (1989). Another useful vector herein is plasmid pBI101.2 that is available from Clontech Laboratories, Inc. (Palo Alto, CA).

A variety of plant viruses that can be employed as vectors are known in the art and include cauliflower mosaic virus (CaMV), geminivirus, brome mosaic virus, and tobacco mosaic virus.

A polynucleotide of the present invention can be expressed in either sense or anti-sense orientation as desired. In plant cells, it has been shown that antisense RNA inhibits gene expression by preventing the accumulation of mRNA which encodes the enzyme of interest, see, e.g., Sheehy *et al.*, *Proc. Nat'l. Acad. Sci. (USA)* 85: 8805-8809 (1988); and Hiatt *et al.*, U.S. Patent No. 4,801,340.

Another method of suppression is sense suppression. Introduction of nucleic acid configured in the sense orientation has been shown to be an effective means by which to block the transcription of target genes. For an example of the use of this method to modulate expression of endogenous genes see, Napoli *et al.*, *The Plant Cell* 2: 279-289 (1990) and U.S. Patent No. 5,034,323.

A method of down-regulation of the protein involves using PEST sequences that provide a target for degradation of the protein.

Catalytic RNA molecules or ribozymes can also be used to inhibit expression of plant genes. The inclusion of ribozyme sequences within antisense RNAs confers RNA-cleaving activity upon them, thereby increasing the activity of the constructs. The design and use of target RNA-specific ribozymes is described in Haseloff *et al.*, *Nature* 334: 585-591 (1988).

A variety of cross-linking agents, alkylating agents and radical generating species as pendant groups on polynucleotides of the present invention can be used to bind, label, detect, and/or cleave nucleic acids. For example, Vlassov, V. V., *et al., Nucleic Acids Res* (1986) 14:4065-4076, describe covalent bonding of a single-stranded DNA fragment with alkylating derivatives of nucleotides complementary to target sequences. A report of similar work by the same group is that by Knorre, D. G., *et al., Biochimie* (1985) 67:785-789. Iverson and Dervan also showed sequence-specific cleavage of single-stranded DNA mediated by incorporation of a modified nucleotide which was capable of activating cleavage (*J Am Chem Soc* (1987) 109:1241-1243). Meyer, R. B., *et al., J Am Chem Soc* (1989) 111:8517-8519, effect covalent crosslinking to a target nucleotide using an alkylating agent complementary to the single-stranded target nucleotide sequence. A photoactivated crosslinking to single-stranded oligonucleotides mediated by psoralen was disclosed by Lee, B. L., *et al., Biochemistry* (1988) 27:3197-3203. Use of crosslinking in triple-helix forming probes was also disclosed by Home, *et al., J Am Chem Soc* (1990) 112:2435-2437. Use of N4, N4-ethanocytosine as an alkylating agent to crosslink to single-stranded oligonucleotides has also been described by Webb and Matteucci, *J Am Chem Soc* (1986) 108:2764-2765; *Nucleic Acids Res* (1986) 14:7661-7674; Feteritz *et al., J. Am. Chem. Soc.* 113:4000 (1991). Various compounds to bind, detect, label, and/or cleave nucleic acids are known in the art. See, for example, U.S. Patent Nos. 5,543,507; 5,672,593; 5,484,908; 5,256,648; and, 5,681941.

PROTEINS

Proteins of the present invention include proteins derived from the native protein by deletion (so-called truncation), and addition or substitution of one or more amino acids at one or more sites in the native protein. Such variants may result from, for example, genetic polymorphism or from human manipulation. Methods for such manipulations are generally known in the art.

For example, amino acid sequence variants of the polypeptide can be prepared by mutations in the cloned DNA sequence encoding the native protein of interest. Methods for mutagenesis and nucleotide sequence alterations are well known in the art. See, for example, Walker and Gaastra, eds. (1983) *Techniques in Molecular Biology* (MacMillan Publishing Company, New York); Kunkel (1985)

Proc. Natl. Acad. Sci. USA 82:488-492; Kunkel *et al.* (1987) *Methods Enzymol.* 154:367-382; Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor, New York); U.S. Patent No. 4,873,192; and the references cited therein; herein incorporated by reference. Guidance as to appropriate amino acid substitutions that do not affect biological activity of the protein of interest may be found in the model of Dayhoff *et al.* (1978) *Atlas of Protein Sequence and Structure* (Natl. Biomed. Res. Found., Washington, D.C.), herein incorporated by reference. Conservative substitutions, such as exchanging one amino acid with another having similar properties, may be preferred.

In constructing variants of the proteins of interest, modifications to the nucleotide sequences encoding the variants will be made such that variants continue to possess the desired activity. Obviously, any mutations made in the DNA encoding the variant protein must not place the sequence out of reading frame and preferably will not create complementary regions that could produce secondary mRNA structure. See EP Patent Application Publication No. 75,444.

The isolated proteins of the present invention include a polypeptide comprising at least 23 contiguous amino acids encoded by any one of the nucleic acids of the present invention, or polypeptides which are conservatively modified variants thereof. The proteins of the present invention or variants thereof can comprise any number of contiguous amino acid residues from a polypeptide of the present invention, wherein that number is selected from the group of integers consisting of from 23 to the number of residues in a full-length polypeptide of the present invention. Optionally, this subsequence of contiguous amino acids is at least 25, 30, 35, or 40 amino acids in length, often at least 50, 60, 70, 80, or 90 amino acids in length.

The present invention includes catalytically active polypeptides (i.e., enzymes). Catalytically active polypeptides will generally have a specific activity of at least 20%, 30%, or 40%, and preferably at least 50%, 60%, or 70%, and most preferably at least 80%, 90%, or 95% that of the native (non-synthetic), endogenous polypeptide. Further, the substrate specificity (k_{cat}/K_m) is optionally substantially similar to the native (non-synthetic), endogenous polypeptide. Typically, the K_m will be at least 30%, 40%, or 50%, that of the native (non-synthetic), endogenous polypeptide; and more preferably at least 60%, 70%, 80%,

or 90%. Methods of assaying and quantifying measures of enzymatic activity and substrate specificity (k_{cat}/K_m), are well known to those of skill in the art.

The present invention includes modifications that can be made to an inventive protein without diminishing its biological/catalytic activity. Some modifications may be made to facilitate the cloning, expression, or incorporation of the targeting molecule into a fusion protein. Such modifications are well known to those of skill in the art and include, for example, a methionine added at the amino terminus to provide an initiation site, or additional amino acids (*e.g.*, poly His) placed on either terminus to create conveniently located restriction sites or termination codons or purification sequences.

A protein of the present invention can be expressed in a recombinantly engineered cell such as bacteria, yeast, insect, mammalian, or preferably plant cells. The cells produce the protein in a non-natural condition (e.g., in quantity, composition, location, and/or time), because they have been genetically altered through human intervention to do so.

Typically, an intermediate host cell will be used in the practice of this invention to increase the copy number of the cloning vector. With an increased copy number, the vector containing the nucleic acid of interest can be isolated in significant quantities for introduction into the desired plant cells.

Host cells that can be used in the practice of this invention include prokaryotes, including bacterial hosts such as *Eschericia coli*, *Salmonella typhimurium*, and *Serratia marcescens*. Eukaryotic hosts such as yeast or filamentous fungi may also be used in this invention. It preferred to use plant promoters that do not cause expression of the polypeptide in bacteria.

Commonly used prokaryotic control sequences include promoters such as the beta lactamase (penicillinase) and lactose (lac) promoter systems (Chang *et al.*, Nature 198:1056 (1977)), the tryptophan (trp) promoter system (Goeddel *et al.*, Nucleic Acids Res. 8:4057 (1980)) and the lambda derived P L promoter and N-gene ribosome binding site (Shimatake *et al.*, Nature 292:128 (1981)). The inclusion of selection markers in DNA vectors transfected in *E. coli* is also useful. Examples of such markers include genes specifying resistance to ampicillin, tetracycline, or chloramphenicol.

The vector is selected to allow introduction into the appropriate host cell. Bacterial vectors are typically of plasmid or phage origin. Expression systems for

expressing a protein of the present invention are available using *Bacillus sp.* and *Salmonella* (Palva, *et al.*, *Gene* 22: 229-235 (1983); Mosbach, *et al.*, *Nature* 302: 543-545 (1983)).

Synthesis of heterologous proteins in yeast is well known. See Sherman, F., *et al.*, *Methods in Yeast Genetics*, Cold Spring Harbor Laboratory (1982). Two widely utilized yeast for production of eukaryotic proteins are *Saccharomyces cerevisiae* and *Pichia pastoris*. Vectors, strains, and protocols for expression in *Saccharomyces* and *Pichia* are known in the art and available from commercial suppliers (e.g., Invitrogen). Suitable vectors usually have expression control sequences, such as promoters, including 3-phosphoglycerate kinase or alcohol oxidase, and an origin of replication, termination sequences and the like as desired.

A protein of the present invention, once expressed, can be isolated from yeast by lysing the cells and applying standard protein isolation techniques to the lysates. The monitoring of the purification process can be accomplished by using Western blot techniques or radioimmunoassay of other standard immunoassay techniques.

The proteins of the present invention can also be constructed using non-cellular synthetic methods. Solid phase synthesis of proteins of less than about 50 amino acids in length may be accomplished by attaching the C-terminal amino acid of the sequence to an insoluble support followed by sequential addition of the remaining amino acids in the sequence. Techniques for solid phase synthesis are described by Barany and Merrifield, Solid-Phase Peptide Synthesis, pp. 3-284 in *The Peptides: Analysis, Synthesis, Biology. Vol. 2: Special Methods in Peptide Synthesis, Part A.*; Merrifield, *et al.*, *J. Am. Chem. Soc.* 85: 2149-2156 (1963), and *Stewart et al., Solid Phase Peptide Synthesis, 2nd ed.,* Pierce Chem. Co., Rockford, Ill. (1984). Proteins of greater length may be synthesized by condensation of the amino and carboxy termini of shorter fragments. Methods of forming peptide bonds by activation of a carboxy terminal end (e.g., by the use of the coupling reagent N,N'-dicycylohexylcarbodiimide) is known to those of skill.

The proteins of this invention may be purified to substantial purity by standard techniques well known in the art, including detergent solubilization, selective precipitation with such substances as ammonium sulfate, column chromatography, immunopurification methods, and others. *See*, for instance, R.

Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag: New York (1982); Deutscher, *Guide to Protein Purification*, Academic Press (1990). For example, antibodies may be raised to the proteins as described herein. Purification from *E. coli* can be achieved following procedures described in U.S. Patent No. 4,511,503. Detection of the expressed protein is achieved by methods known in the art and include, for example, radioimmunoassays, Western blotting techniques or immunoprecipitation.

The present invention further provides a method for modulating (i.e., increasing or decreasing) the concentration or composition of the polypeptides of the present invention in a plant or part thereof. Modulation of the polypeptides can be effected by increasing or decreasing the concentration and/or the composition of the polypeptides in a plant. The method comprises transforming a plant cell with an expression cassette comprising a polynucleotide of the present invention to obtain a transformed plant cell, growing the transformed plant cell under plant forming conditions, and inducing expression of the polynucleotide in the plant for a time sufficient to modulate concentration and/or composition of the polypeptides in the plant or plant part.

In some embodiments, the content and/or composition of polypeptides of the present invention in a plant may be modulated by altering, *in vivo* or *in vitro*, the promoter of a non-isolated gene of the present invention to up- or down-regulate gene expression. In some embodiments, the coding regions of native genes of the present invention can be altered via substitution, addition, insertion, or deletion to decrease activity of the encoded enzyme. See, e.g., Kmiec, U.S. Patent 5,565,350; Zarling *et al.*, PCT/US93/03868.

In particular, modulating cell cycle proteins are expected to provide a positive growth advantage and increase crop yield. Cell cycle nucleic acids can be adducted to a second nucleic acid sequence encoding a DNA-binding domain, for use in two-hybrid systems to identify DP-interacting proteins. It is expected that modulating the level of cell cycle protein, i.e. overexpression of *DP* will increase endoreduplication. Endoreduplication is expected to increase the size of the seed, the size of the endosperm and the amount of protein in the seed.

Also, modulating cell cycle proteins affects the cell number in a tissue of a plant, thereby affecting the size and characteristics of that tissue organ. Modulation could affect any plant tissue such as, but not limited to root, seed,

tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leave, stem, seed, trunk, meristem or fruit. Changes in plant tissue will influence quality traits, agronomic traits and susceptibility to disease and insects.

In some embodiments, an isolated nucleic acid (e.g., a vector) comprising a promoter sequence is transfected into a plant cell. Subsequently, a plant cell comprising the isolated nucleic acid is selected for by means known to those of skill in the art such as, but not limited to, Southern blot, DNA sequencing, or PCR analysis using primers specific to the promoter and to the nucleic acid and detecting amplicons produced therefrom. A plant or plant part altered or modified by the foregoing embodiments is grown under plant forming conditions for a time sufficient to modulate the concentration and/or composition of polypeptides of the present invention in the plant. Plant forming conditions are well known in the art.

In general, concentration of the polypeptides is increased or decreased by at least 2%, 3%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% relative to a native control plant, plant part, or cell lacking the aforementioned expression cassette. Modulation in the present invention may occur during and/or subsequent to growth of the plant to the desired stage of development.

Modulating nucleic acid expression temporally and/or in particular tissues can be controlled by employing the appropriate promoter operably linked to a polynucleotide of the present invention in, for example, sense or antisense orientation as discussed in greater detail above. Induction of expression of a polynucleotide of the present invention can also be controlled by exogenous administration of an effective amount of inducing compound. Inducible promoters and inducing compounds that activate expression from these promoters are well known in the art.

In preferred embodiments, the polypeptides of the present invention are modulated in monocots, preferably cereals, or dicots. Preferred plants include corn, soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice, barley or millet. In another embodiment the polypeptides of this present invention are modulated in bacteria and yeast.

Means of detecting the proteins of the present invention are not critical aspects of the present invention. In a preferred embodiment, the proteins are detected and/or quantified using any of a number of well recognized immunological binding assays (*see, e.g.,* U.S. Patents 4,366,241; 4,376,110;

4,517,288; and 4,837,168). For a review of the general immunoassays, see also Methods in Cell Biology, Vol. 37: *Antibodies in Cell Biology*, Asai, Ed., Academic Press, Inc. New York (1993); *Basic and Clinical Immunology* 7th Edition, Stites & Terr, Eds. (1991). Moreover, the immunoassays of the present invention can be performed in any of several configurations, *e.g.*, those reviewed in *Enzyme Immunoassay*, Maggio, Ed., CRC Press, Boca Raton, Florida (1980); Tijan, Practice and Theory of Enzyme Immunoassays, *Laboratory Techniques in Biochemistry and Molecular Biology*, Elsevier Science Publishers B.V., Amsterdam (1985); Harlow and Lane, *supra*; *Immunoassay: A Practical Guide*, Chan, Ed., Academic Press, Orlando, FL (1987); *Principles and Practice of Immunoassays*, Price and Newman Eds., Stockton Press, NY (1991); and *Non-isotopic Immunoassays*, Ngo, Ed., Plenum Press, NY (1988).

Typical methods for detecting proteins include Western blot (immunoblot) analysis, analytic biochemical methods such as electrophoresis, capillary electrophoresis, high performance liquid chromatography (HPLC), thin layer chromatography (TLC), hyperdiffusion chromatography, and the like, and various immunological methods such as fluid or gel precipitin reactions, immunodiffusion (single or double), immunoelectrophoresis, radioimmunoassays (RIAs), enzyme-linked immunosorbent assays (ELISAs), immunofluorescent assays, and the like.

Non-radioactive labels are often attached by indirect means. Generally, a ligand molecule (*e.g.*, biotin) is covalently bound to the molecule. The ligand then binds to an anti-ligand (*e.g.*, streptavidin) molecule which is either inherently detectable or covalently bound to a signal system, such as a detectable enzyme, a fluorescent compound, or a chemiluminescent compound. A number of ligands and anti-ligands can be used. Where a ligand has a natural anti-ligand, for example, biotin, thyroxine, and cortisol, it can be used in conjunction with the labeled, naturally occurring anti-ligands. Alternatively, any haptenic or antigenic compound can be used in combination with an antibody.

The molecules can also be conjugated directly to signal generating compounds, *e.g.*, by conjugation with an enzyme or fluorophore. Enzymes of interest as labels will primarily be hydrolases, particularly phosphatases, esterases and glycosidases, or oxidoreductases, particularly peroxidases. Fluorescent compounds include fluorescein and its derivatives, rhodamine and its derivatives, dansyl, umbelliferone, etc. Chemiluminescent compounds include luciferin, and

2,3-dihydrophthalazinediones, e.g., luminol. For a review of various labeling or signal producing systems which may be used, see, U.S. Patent No. 4,391,904, which is incorporated herein by reference.

Some assay formats do not require the use of labeled components. For instance, agglutination assays can be used to detect the presence of the target antibodies. In this case, antigen-coated particles are agglutinated by samples comprising the target antibodies. In this format, none of the components need be labeled and the presence of the target antibody is detected by simple visual inspection.

The proteins of the present invention can be used for identifying compounds that bind to (e.g., substrates), and/or increase or decrease (i.e., modulate) the enzymatic activity of, catalytically active polypeptides of the present invention. The method comprises contacting a polypeptide of the present invention with a compound whose ability to bind to or modulate enzyme activity is to be determined. The polypeptide employed will have at least 20%, preferably at least 30% or 40%, more preferably at least 50% or 60%, and most preferably at least 70% or 80% of the specific activity of the native, full-length polypeptide of the present invention (e.g., enzyme). Methods of measuring enzyme kinetics are well known in the art. See, e.g., Segel, *Biochemical Calculations*, 2nd ed., John Wiley and Sons, New York (1976).

Antibodies can be raised to a protein of the present invention, including individual, allelic, strain, or species variants, and fragments thereof, both in their naturally occurring (full-length) forms and in recombinant forms. Additionally, antibodies are raised to these proteins in either their native configurations or in non-native configurations. Anti-idiotypic antibodies can also be generated. Many methods of making antibodies are known to persons of skill.

In some instances, it is desirable to prepare monoclonal antibodies from various mammalian hosts, such as mice, rodents, primates, humans, *etc.* Description of techniques for preparing such monoclonal antibodies are found in, *e.g.*, *Basic and Clinical Immunology,* 4th ed., Stites *et al.*, Eds., Lange Medical Publications, Los Altos, CA, and references cited therein; Harlow and Lane, *Supra*; Goding, *Monoclonal Antibodies: Principles and Practice*, 2nd ed., Academic Press, New York, NY (1986); and Kohler and Milstein, *Nature* 256: 495-497 (1975).

Other suitable techniques involve selection of libraries of recombinant antibodies in phage or similar vectors (*see, e.g.,* Huse *et al.*, *Science* 246: 1275-1281 (1989); and Ward, *et al.*, *Nature* 341: 544-546 (1989); and Vaughan *et al.*, *Nature Biotechnology*, 14: 309-314 (1996)). Alternatively, high avidity human monoclonal antibodies can be obtained from transgenic mice comprising fragments of the unrearranged human heavy and light chain Ig loci (i.e., minilocus transgenic mice). Fishwild *et al.*, *Nature Biotech.*, 14: 845-851 (1996). Also, recombinant immunoglobulins may be produced. See, Cabilly, U.S. Patent No. 4,816,567; and Queen *et al.*, *Proc. Nat'l Acad. Sci.* 86: 10029-10033 (1989).

The antibodies of this invention can be used for affinity chromatography in isolating proteins of the present invention, for screening expression libraries for particular expression products such as normal or abnormal protein or for raising anti-idiotypic antibodies which are useful for detecting or diagnosing various pathological conditions related to the presence of the respective antigens.

Frequently, the proteins and antibodies of the present invention will be labeled by joining, either covalently or non-covalently, a substance which provides for a detectable signal. A wide variety of labels and conjugation techniques are known and are reported extensively in both the scientific and patent literature. Suitable labels include radionucleotides, enzymes, substrates, cofactors, inhibitors, fluorescent moieties, chemiluminescent moieties, magnetic particles, and the like.

Transfection/Transformation of Cells

The method of transformation/transfection is not critical to the invention; various methods of transformation or transfection are currently available. As newer methods are available to transform crops or other host cells they may be directly applied. Accordingly, a wide variety of methods have been developed to insert a DNA sequence into the genome of a host cell to obtain the transcription and/or translation of the sequence to effect phenotypic changes in the organism. Thus, any method that provides for efficient transformation/transfection may be employed.

A DNA sequence coding for the desired polynucleotide of the present invention, for example a cDNA, RNA or a genomic sequence, will be used to construct an expression cassette that can be introduced into the desired plant.

Isolated nucleic acid acids of the present invention can be introduced into plants according techniques known in the art. Generally, expression cassettes as described above and suitable for transformation of plant cells are prepared.

Techniques for transforming a wide variety of higher plant species are well known and described in the technical, scientific, and patent literature. See, for example, Weising *et al., Ann. Rev. Genet.* 22: 421-477 (1988). For example, the DNA construct may be introduced directly into the genomic DNA of the plant cell using techniques such as electroporation, PEG-mediated transfection, particle bombardment, silicon fiber delivery, or microinjection of plant cell protoplasts or embryogenic callus. See, e.g., Tomes, *et al.,* Direct DNA Transfer into Intact Plant Cells Via Microprojectile Bombardment. pp.197-213 in Plant Cell, Tissue and Organ Culture, Fundamental Methods. eds. O. L. Gamborg and G.C. Phillips. Springer-Verlag Berlin Heidelberg New York, 1995. Alternatively, the DNA constructs may be combined with suitable T-DNA flanking regions and introduced into a conventional *Agrobacterium tumefaciens* host vector. The virulence functions of the *Agrobacteriurn tumefaciens* host will direct the insertion of the construct and adjacent marker into the plant cell DNA when the cell is infected by the bacteria. See, U.S. Patent No. 5,591,616.

The introduction of DNA constructs using polyethylene glycol precipitation is described in Paszkowski *et al., Embo J.* 3: 2717-2722 (1984). Electroporation techniques are described in Fromm *et al., Proc. Natl. Acad. Sci.* 82: 5824 (1985). Ballistic transformation techniques are described in Klein *et al., Nature* 327: 70-73 (1987).

Agrobacterium tumefaciens-meditated transformation techniques are well described in the scientific literature. See, for example Horsch *et al., Science* 233: 496-498 (1984), and Fraley *et al., Proc. Natl. Acad. Sci.* 80: 4803 (1983). For instance, *Agrobacterium* transformation of maize is described in U.S. Patent Nos. 5,550,318 and 5,981,840, which are incorporated by reference herewithin.

Other methods of transfection or transformation include (1) *Agrobacterium rhizogenes*-mediated transformation (see, e.g., Lichtenstein and Fuller In: Genetic Engineering, vol. 6, PWJ Rigby, Ed., London, Academic Press, 1987; and Lichtenstein, C. P., and Draper, J,. In: DNA Cloning, Vol. II, D. M. Glover, Ed., Oxford, IRI Press, 1985),Application PCT/US87/02512 (WO 88/02405 published Apr. 7, 1988) describes the use of *A. rhizogenes* strain A4 and its Ri plasmid along

with *A. tumefaciens* vectors pARC8 or pARC16 (2) liposome-mediated DNA uptake (see, e.g., Freeman *et al.*, Plant Cell Physiol. 25: 1353, 1984), (3) the vortexing method (see, e.g., Kindle, *Proc. Natl. Acad. Sci.*, USA 87: 1228, (1990).

DNA can also be introduced into plants by direct DNA transfer into pollen as described by Zhou *et al.*, Methods in Enzymology, 101:433 (1983); D. Hess, Intern Rev. Cytol., 107:367 (1987); Luo *et al.*, Plane Mol. Biol. Reporter, 6:165 (1988). Expression of polypeptide coding nucleic acids can be obtained by injection of the DNA into reproductive organs of a plant as described by Pena *et al.*, Nature, 325.:274 (1987). DNA can also be injected directly into the cells of immature embryos and the rehydration of desiccated embryos as described by Neuhaus *et al.*, Theor. Appl. Genet., 75:30 (1987); and Benbrook *et al.*, in Proceedings Bio Expo 1986, Butterworth, Stoneham, Mass., pp. 27-54 (1986).

Animal and lower eukaryotic (e.g., yeast) host cells are competent or rendered competent for transfection by various means. There are several well-known methods of introducing DNA into animal cells. These include: calcium phosphate precipitation, fusion of the recipient cells with bacterial protoplasts containing the DNA, treatment of the recipient cells with liposomes containing the DNA, DEAE dextran, electroporation, biolistics, and micro-injection of the DNA directly into the cells. The transfected cells are cultured by means well known in the art. Kuchler, R.J., *Biochemical Methods in Cell Culture and Virology*, Dowden, Hutchinson and Ross, Inc. (1977).

Transgenic Plant Regeneration

Transformed plant cells which are derived by any of the above transformation techniques can be cultured to regenerate a whole plant which possesses the transformed genotype. Such regeneration techniques often rely on manipulation of certain phytohormones in a tissue culture growth medium, typically relying on a biocide and/or herbicide marker which has been introduced together with a polynucleotide of the present invention. For transformation and regeneration of maize see, Gordon-Kamm *et al., The Plant Cell*, 2:603-618 (1990).

Plants cells transformed with a plant expression vector can be regenerated, e.g., from single cells, callus tissue or leaf discs according to standard plant tissue culture techniques. It is well known in the art that various cells, tissues, and organs from almost any plant can be successfully cultured to regenerate an entire

plant. Plant regeneration from cultured protoplasts is described in Evans *et al.*, *Protoplasts Isolation and Culture*, *Handbook of Plant Cell Culture*, Macmillan Publishing Company, New York, pp. 124-176 (1983); and Binding, *Regeneration of Plants*, *Plant Protoplasts*, CRC Press, Boca Raton, pp. 21-73 (1985).

The regeneration of plants containing the foreign gene introduced by *Agrobacterium* can be achieved as described by Horsch *et al.*, *Science*, 227:1229-1231 (1985) and Fraley *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, 80:4803 (1983). This procedure typically produces shoots within two to four weeks and these transformant shoots are then transferred to an appropriate root-inducing medium containing the selective agent and an antibiotic to prevent bacterial growth. Transgenic plants of the present invention may be fertile or sterile.

Regeneration can also be obtained from plant callus, explants, organs, or parts thereof. Such regeneration techniques are described generally in Klee *et al.*, *Ann. Rev. of Plant Phys.* 38: 467-486 (1987). The regeneration of plants from either single plant protoplasts or various explants is well known in the art. See, for example, Methods for Plant Molecular Biology, A. Weissbach and H. Weissbach, eds., Academic Press, Inc., San Diego, Calif. (1988). For maize cell culture and regeneration see generally, *The Maize Handbook*, Freeling and Walbot, Eds., Springer, New York (1994); *Corn and Corn Improvement*, 3rd edition, Sprague and Dudley Eds., American Society of Agronomy, Madison, Wisconsin (1988).

One of skill will recognize that after the expression cassette is stably incorporated in transgenic plants and confirmed to be operable, it can be introduced into other plants by sexual crossing. Any of a number of standard breeding techniques can be used, depending upon the species to be crossed.

In vegetatively propagated crops, mature transgenic plants can be propagated by the taking of cuttings or by tissue culture techniques to produce multiple identical plants. Selection of desirable transgenics is made and new varieties are obtained and propagated vegetatively for commercial use. In seed propagated crops, mature transgenic plants can be self crossed to produce a homozygous inbred plant. The inbred plant produces seed containing the newly introduced heterologous nucleic acid. These seeds can be grown to produce plants that would produce the selected phenotype.

Parts obtained from the regenerated plant, such as flowers, seeds, leaves, branches, fruit, and the like are included in the invention, provided that these parts

comprise cells comprising the isolated nucleic acid of the present invention. Progeny and variants, and mutants of the regenerated plants are also included within the scope of the invention, provided that these parts comprise the introduced nucleic acid sequences.

Transgenic plants expressing a selectable marker can be screened for transmission of the nucleic acid of the present invention by, for example, standard immunoblot and DNA detection techniques. Transgenic lines are also typically evaluated on levels of expression of the heterologous nucleic acid. Expression at the RNA level can be determined initially to identify and quantitate expression-positive plants. Standard techniques for RNA analysis can be employed and include PCR amplification assays using oligonucleotide primers designed to amplify only the heterologous RNA templates and solution hybridization assays using heterologous nucleic acid-specific probes. The RNA-positive plants can then analyzed for protein expression by Western immunoblot analysis using the specifically reactive antibodies of the present invention. In addition, *in situ* hybridization and immunocytochemistry according to standard protocols can be done using heterologous nucleic acid specific polynucleotide probes and antibodies, respectively, to localize sites of expression within transgenic tissue. Generally, a number of transgenic lines are usually screened for the incorporated nucleic acid to identify and select plants with the most appropriate expression profiles.

A preferred embodiment is a transgenic plant that is homozygous for the added heterologous nucleic acid; i.e., a transgenic plant that contains two added nucleic acid sequences, one gene at the same locus on each chromosome of a chromosome pair. A homozygous transgenic plant can be obtained by sexually mating (selfing) a heterozygous transgenic plant that contains a single added heterologous nucleic acid, germinating some of the seed produced and analyzing the resulting plants produced for altered expression of a polynucleotide of the present invention relative to a control plant (i.e., native, non-transgenic). Back-crossing to a parental plant and out-crossing with a non- transgenic plant are also contemplated.

Modulating Cell Cycle Protein Content and/or Composition

The present invention further provides a method for modulating (i.e., increasing or decreasing) cell cycle protein content or composition in a plant or part thereof. Modulation can be effected by increasing or decreasing the cell cycle protein content (i.e., the total amount of cell cycle protein) and/or the cell cycle protein composition (the ratio of various cell cycle monomers in the plant) in a plant. The method comprises transforming a plant cell, transiently or stably, with a recombinant expression cassette comprising a polynucleotide of the present invention as described above to obtain a transformed plant cell.

For stably transformed plant cells, the method comprises growing the transformed plant cell under plant forming conditions, and expressing the polynucleotide of the present invention in the plant for a time sufficient to modulate cell cycle protein content and/or composition in the plant or plant part.

In some embodiments, plant cell division may be modulated by altering, *in vivo* or *in vitro*, the promoter of a non-isolated cell cycle gene to up- or down-regulate gene expression. In some embodiments, the coding regions of native cell cycle genes can be altered via substitution, addition, insertion, or deletion to decrease activity of the encoded enzyme. See, e.g., Kmiec, U.S. Patent 5,565,350, Zarling *et al.*, PCT/US93/03868. And in some embodiments, an isolated nucleic acid (e.g., a vector) comprising a promoter sequence is transfected into a plant cell. Subsequently, a plant cell comprising the promoter operably linked to a polynucleotide of the present invention is selected for by means known to those of skill in the art such as, but not limited to, Southern blot, DNA sequencing, or PCR analysis using primers specific to the promoter and to the gene and detecting amplicons produced therefrom. A plant or plant part altered or modified by the foregoing embodiments is grown under plant forming conditions for a time sufficient to modulate cell cycle protein content and/or composition in the plant. Plant forming conditions are well known in the art and discussed briefly, *supra*.

In general, content or composition is increased or decreased by at least 2%, 3%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% relative to a native control plant, plant part, or cell lacking the aforementioned recombinant expression cassette. Modulation in the present invention may occur during and/or subsequent to growth of the plant to the desired stage of development.

Modulating nucleic acid expression temporally and/or in particular tissues can be controlled by employing the appropriate promoter operably linked to a polynucleotide of the present invention in, for example, sense or antisense orientation as discussed in greater detail, *supra*. Induction of expression of a polynucleotide of the present invention can also be controlled by exogenous administration of an effective amount of inducing compound. Inducible promoters and inducing compounds which activate expression from these promoters are well known in the art. In preferred embodiments, cell division is modulated in monocots, particularly maize.

UTR's and Codon Preference

In general, translational efficiency has been found to be regulated by specific sequence elements in the 5' non-coding or untranslated region (5' UTR) of the RNA. Positive sequence motifs include translational initiation consensus sequences (Kozak, Nucleic Acids Res.15:8125 (1987)) and the 5<G> 7 methyl GpppG cap structure (Drummond *et al.*, Nucleic Acids Res. 13:7375 (1985)). Negative elements include stable intramolecular 5' UTR stem-loop structures (Muesing *et al.*, Cell 48:691 (1987)) and AUG sequences or short open reading frames preceded by an appropriate AUG in the 5' UTR (Kozak, *supra,* Rao *et al.*, Mol. and Cell. Biol. 8:284 (1988)). Accordingly, the present invention provides 5' and/or 3' UTR regions for modulation of translation of heterologous coding sequences.

Further, the polypeptide-encoding segments of the polynucleotides of the present invention can be modified to alter codon usage. Altered codon usage can be employed to alter translational efficiency and/or to optimize the coding sequence for expression in a desired host or to optimize the codon usage in a heterologous sequence for expression in maize. Codon usage in the coding regions of the polynucleotides of the present invention can be analyzed statistically using commercially available software packages such as "Codon Preference" available from the University of Wisconsin Genetics Computer Group (see Devereaux *et al.*, Nucleic Acids Res. 12:387-395 (1984)) or MacVector 4.1 (Eastman Kodak Co., New Haven, Conn.). Thus, the present invention provides a codon usage frequency characteristic of the coding region of at least one of the polynucleotides of the present invention. The number of polynucleotides that can

be used to determine a codon usage frequency can be any integer from 1 to the number of polynucleotides of the present invention as provided herein. Optionally, the polynucleotides will be full-length sequences. An exemplary number of sequences for statistical analysis can be at least 1, 5, 10, 20, 50, or 100.

Sequence Shuffling

The present invention provides methods for sequence shuffling using polynucleotides of the present invention, and compositions resulting therefrom. Sequence shuffling is described in PCT publication No. 96/19256. See also, Zhang, J. H., *et al.* Proc. Natl. Acad. Sci. USA 94:4504-4509 (1997). Generally, sequence shuffling provides a means for generating libraries of polynucleotides having a desired characteristic which can be selected or screened for. Libraries of recombinant polynucleotides are generated from a population of related sequence polynucleotides which comprise sequence regions which have substantial sequence identity and can be homologously recombined in vitro or in vivo. The population of sequence-recombined polynucleotides comprises a subpopulation of polynucleotides which possess desired or advantageous characteristics and which can be selected by a suitable selection or screening method. The characteristics can be any property or attribute capable of being selected for or detected in a screening system, and may include properties of: an encoded protein, a transcriptional element, a sequence controlling transcription, RNA processing, RNA stability, chromatin conformation, translation, or other expression property of a gene or transgene, a replicative element, a protein-binding element, or the like, such as any feature which confers a selectable or detectable property. In some embodiments, the selected characteristic will be an increased K_m and/or K_{cat} over the wild-type protein as provided herein. In other embodiments, a protein or polynculeotide generated from sequence shuffling will have a ligand binding affinity greater than the non-shuffled wild-type polynucleotide. The increase in such properties can be at least 110%, 120%, 130%, 140% or at least 150% of the wild-type value.

Detection of Nucleic Acids

The present invention further provides methods for detecting a polynucleotide of the present invention in a nucleic acid sample suspected of

comprising a polynucleotide of the present invention, such as a plant cell lysate, particularly a lysate of corn. In some embodiments, a cell cycle gene or portion thereof can be amplified prior to the step of contacting the nucleic acid sample with a polynucleotide of the present invention. The nucleic acid sample is contacted with the polynucleotide to form a hybridization complex. The polynucleotide hybridizes under stringent conditions to a gene encoding a polypeptide of the present invention. Formation of the hybridization complex is used to detect a gene encoding a polypeptide of the present invention in the nucleic acid sample. Those of skill will appreciate that an isolated nucleic acid comprising a polynucleotide of the present invention should lack cross-hybridizing sequences in common with non-cell cycle genes that would yield a false positive result.

Detection of the hybridization complex can be achieved using any number of well known methods. For example, the nucleic acid sample, or a portion thereof, may be assayed by hybridization formats including but not limited to, solution phase, solid phase, mixed phase, or *in situ* hybridization assays. Briefly, in solution (or liquid) phase hybridizations, both the target nucleic acid and the probe or primer are free to interact in the reaction mixture. In solid phase hybridization assays, probes or primers are typically linked to a solid support where they are available for hybridization with target nucleic in solution. In mixed phase, nucleic acid intermediates in solution hybridize to target nucleic acids in solution as well as to a nucleic acid linked to a solid support. In *in situ* hybridization, the target nucleic acid is liberated from its cellular surroundings in such as to be available for hybridization within the cell while preserving the cellular morphology for subsequent interpretation and analysis. The following articles provide an overview of the various hybridization assay formats: Singer *et al.*, Biotechniques 4(3):230-250 (1986); Haase *et al.*, *Methods in Virology*, Vol. VII, pp. 189-226 (1984); Wilkinson, The theory and practice of in situ hybridization in: *In situ Hybridization*, D.G. Wilkinson, Ed., IRL Press, Oxford University Press, Oxford; and *Nucleic Acid Hybridization: A Practical Approach*, Hames, B.D. and Higgins, S.J., Eds., IRL Press (1987).

Molecular Markers

The present invention provides a method of genotyping a plant comprising a polynucleotide of the present invention. Genotyping provides a means of distinguishing homologs of a chromosome pair and can be used to differentiate segregants in a plant population. Molecular marker methods can be used for phylogenetic studies, characterizing genetic relationships among crop varieties, identifying crosses or somatic hybrids, localizing chromosomal segments affecting monogenic traits, map based cloning, and the study of quantitative inheritance. See, e.g., *Plant Molecular Biology: A Laboratory Manual*, Chapter 7, Clark, Ed., Springer-Verlag, Berlin (1997). For molecular marker methods, see generally, The DNA Revolution by Andrew H. Paterson 1996 (Chapter 2) in: Genome Mapping in Plants (ed. Andrew H. Paterson) by Academic Press/R. G. Landis Company, Austin, Texas, pp.7-21.

The particular method of genotyping in the present invention may employ any number of molecular marker analytic techniques such as, but not limited to, restriction fragment length polymorphisms (RFLPs). RFLPs are the product of allelic differences between DNA restriction fragments caused by nucleotide sequence variability. Thus, the present invention further provides a means to follow segregation of a gene or nucleic acid of the present invention as well as chromosomal sequences genetically linked to these genes or nucleic acids using such techniques as RFLP analysis.

Plants that can be used in the method of the invention include monocotyledonous and dicotyledonous plants. Preferred plants include corn, soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice, barley and millet.

Seeds derived from plants regenerated from transformed plant cells, plant parts or plant tissues, or progeny derived from the regenerated transformed plants, may be used directly as feed or food, or further processing may occur.

Expression of the inventive nucleic acids in plants, such as corn is expected to enhance growth and biomass accumulation. Other more specialized applications exist for these nucleic acids at the whole plant level. It has been demonstrated that endoreduplication occurs in numerous cell types within plants, but this is particularly prevalent in maize endosperm, the primary seed storage tissue. Under the direction of endosperm-specific promoters, expression of cell

cycle genes (and possibly expression of such genes in conjunction with genes that inhibit mitosis) will further stimulate the process of endoreduplication.

The E2F/Dp heterodimer activates transcription of numerous genes involved in DNA replication, thus playing a role in the G1 to S transition in the cell cycle. In addition, E2F/Dp participates in multiprotein complexes with the retinoblastoma susceptibility protein (Rb), p107 (a homologue of Rb), cyclins and cdk2 (Bagehi *et al.*, 1991; Pandara and La Thangue, 1991; Chelloppan *et al.*, 1991; Chittenden *et al.*, 1991). Binding of Dp to E2F has been found to enhance E2F's ability to bind both DNA and pRb, as well as its ability to regulate (either positively or negatively) the transcription of target genes (Helin *et al.*, 1993; Zhang and Chellappan, 1995; Bandara *et al.*, 1993; Cao *et al.*, 1992).

The DP family of genes have been cloned and well-characterized in Drosophila (Dynlacht *et al.*, 1994, PNAS 91:6359-6363) and in mammalian cells (Girling *et al.*, 1993. Nature 365:468; Helin *et al.*, 1993, Genes Dev. 7:1850-61). A DP homologue from maize (ZmDP) has been cloned and sequenced (see WO 99/53075, incorporated by reference) whose sequence resembles DP protein from other eukaryotes. Utility is demonstrated for use of DP to stimulate DNA replication and to improve transformation.

The present invention will be further described by reference to the following detailed examples. It is understood, however, that there are many extensions, variations, and modifications on the basic theme of the present invention beyond that shown in the examples and description, which are within the spirit and scope of the present invention. All publications, patents, and patent applications cited herein are hereby incorporated by reference.

EXAMPLES

The present invention is further defined in the following Examples, in which all parts and percentages are by weight and degrees are Celsius, unless otherwise stated. It should be understood that these Examples, while indicating preferred embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions.

Example 1: Composition of cDNA libraries: Isolation and Sequencing of cDNA Clones

cDNA libraries representing mRNAs from various maize tissues were prepared. cDNA libraries were prepared in Uni-ZAP tm vectors according to the manufacturer's protocol (Stratagene Cloning Systems, La Jolla, CA). Conversion of the Uni-ZAP XR libraries was accomplished according to the protocol provided by Stratagene. Upon conversion, cDNA inserts were contained in the plasmid vector pBluescript. cDNA inserts from randomly picked bacterial colonies containing recombinant pBluescript plasmids were amplified via polymerase chain reaction using primers specific for vector sequences flanking the inserted cDNA sequences or plasmid DNA was prepared from cultured bacterial cells. Amplified insert DNAs or plasmid DNA were sequenced in dye-primer sequencing reactions to generate partial cDNA sequences (expressed sequence tags or "ESTs", see Adams, M.D. *et al.*, (1991) Science 252: 1651). The resulting ESTs were analyzed using a Perkin Elmer Model 377 fluorescent sequencer.

Identification of cDNA Clones

ESTs encoding were DP identified by conducting BLAST (Basic Local Alignment Search Tool; Altschul, S. F., et a'., (1993) J. Mol. Biol. 215:403410; see also www.ncbi.nlm.nih.gov/BLAST/) searches for similarity to sequences contained in the BLAST "nr" database (comprising all nonredundant GenBank CDS translations, sequences derived from the 3-dimensional structure Brookhaven Protein Data Bann:, the last major release of the SWISS-PROT protein sequence database, EMBL, and DDBJ databases). The cDNA sequences obtained in Example 1 were analyzed for similarity to all publicly available DNA sequences contained in the "nr,' database using the BLASTN algorithm provided by the National Center for Biotechnology Information NCBI). The DNA sequences were translated in all reading frames and compared for similarity to all publicly available protein sequences contained in the "nr" database using the BLASTX algorithm (Gish, W. and States, D. J. (1993) *Nature Genetics 3:266-272)* provided by the NCBI. For convenience, the P-value (probability) of observing a match of a cDNA sequence to a sequence contained in the searched databases merely by chance as calculated by BLAST are reported herein as "pLog" values, which

represent the negative of the logarithm of the reported P-value. Accordingly, the greater the pLog value, the greater the likelihood that the cDNA sequence and the BLAST "hit" represent homologous proteins.

Example 2: Using DP's in a two-hybrid system to identify maize Cell Cycle Genes

The protein encoded by the DP gene plays a prominent role in progression through the cell cycle during the G1→ S transition and S-phase. The proteins encoded by the DP gene family are an important part of the heteroduplex that transactivates many replication and DNA repair genes. DP is an important component in starting the cascade of events leading to DNA replication. Activity of the DP/E2F heteroduplex is further modulated through its association with such proteins as Cyclin A and Cdk2. As such, the DP genes and their encoded proteins can be used to identify other cell cycle regulatory proteins. This can be done using the DP gene as bait (the target fused to the DNA-binding domain) in a yeast two-hybrid screen.

Methods for two-hybrid library construction, cloning of the reporter gene, cloning of the DNA-binding and activation domain hybrid gene cassettes, yeast culture, and transformation of the yeast are all done according to well-established methods (see Sambrook *et al.*, 1990; Ausubel *et al.*, 1990; Hannon and Bartels, 1995). Using this method, *Zea mays* E2F, Cdk2, cycA, Rb and similar genes are identified as components of the activation domain hybrid, and are confirmed through further sequence analysis. Similarly, inhibitors of the E2F/DP complex are identified.

Example 3: DP -bound affinity columns for identifying Dp-binding proteins and their encoding genes

Purified recombinant DP protein can be immobilized on a matrix via a covalent crosslinking or affinity purification as described supra. This matrix can then be used to pull-down proteins that interact with DP proteins. The presence and level of DP interacting proteins can be determined on the basis of immunological assay, activity quantification, SDS-PAGE analysis and other methods. These measures can then be correlated with the reproductive state,

capacity for division, developmental stage, or the quality of different samples. Proteins such as E2F should be identified.

Example 4: Transient DP expression stimulates DNA replication and enhances transgene integration

Regardless of the method of DNA delivery, cells competent for the integration of foreign DNA are actively dividing. There is a growing body of evidence suggesting that integration of foreign DNA occurs in dividing cells (this includes both *Agrobacterium* and direct DNA delivery methods). It has long been observed that dividing transformed cells represent only a fraction of cells that transiently express a transgene. It is well known (in non-plant systems) that the delivery of damaged DNA, (similar to what we introduce by particle gun delivery methods) induces an immediate cell cycle arrest. This inhibition can be obviated by ectopic transient over-expression of positive cell cycle regulators. Regardless of the mechanism of arrest; i.e. presence of damaged DNA or delivery into a non-cycling differentiated cell, stimulation of the cell cycle will increase integration frequencies. To demonstrate this, the DP gene is cloned into a cassette with a constitutive promoter (i.e. either a strong maize promoter such as the ubiquitin promoter including the first ubiquitin intron, or a weak constitutive promoter such as nos). Delivery of the ZmDP gene in an appropriate plant expression cassette (for example, in a UBI::ZmDP::pinII-containing plasmid) along with UBI::bar::pinII can be accomplished through numerous well-established methods for plant cells, including for example particle bombardment, sonication, PEG treatment or electroporation of protoplasts, electroporation of intact tissue, silica-fiber methods, microinjection or *Agrobacterium*-mediated transformation. Using one of the above methods, DNA is introduced into maize cells capable of growth on suitable maize culture medium. Such competent cells can be from maize suspension culture, callus culture on solid medium, freshly isolated immature embryos or meristem cells. Immature embryos of the Hi-II genotype are used as the target for co-delivery of these two plasmids.

Transient expression of the DP gene overcomes the G1/S checkpoint controls, and increases the proportion of recipient-cells (i.e. into which DNA was introduced) that enter S-phase and initiate DNA Replication. This stimulation through the G1/S transition in cells harboring transgenic plasmid DNA provides an

optimal cellular environment for integration of the introduced genes. Cytological methods can be used to verify increased frequencies of progression through S-phase and mitosis (i.e. for cells in which a visual marker such as GFP was transformed alongside DP, the green fluorescent cells will exhibit a higher mitotic index). Cells in S-phase (undergoing DNA replication) can be monitored by detecting nucleotide analog incorporation. For example, following incubation of cells with bromodeoxyuridine (BrdU) incorporation of this thymidine analog can be detected by methods such as antiBrdU immunocytochemistry or through enhancement of Topro3 fluorescence following BrdU labeling. It is expected that DP expression will increase the proportion of cells incorporating BrdU (i.e. a higher percentage of transformed cells will incorporate BrdU relative to untransformed ceils). Increased DNA synthesis can also be monitored using such methods as flow cytometric analysis of protoplasts (or nuclei), in conjunction with appropriate BrdU-insensitive fluorescent DNA labels such as propidium iodide and DAPI or BrdU-detecting methods described above. For example, tissue is homogenized to release nuclei that are analyzed using the flow cytometric analysis for both green fluorescence (from our accompanying GFP marker) and DNA content. Such flow cytometric analysis can demonstrate that expression of a co-transformed GFP reporter correlates with DP-induced changes in the ratios of cells in G1, S and G2. Similar experiments can be run using the fluorescently labeled anti-BrdU antisera to demonstrate that DP expression increased the percentage of cells in S-phase. Cell cycle stage-specific probes can also be used to monitor cell cycle progression. For example, numerous spindle-associated proteins are expressed during a fairly narrow window during mitosis, and antibodies or nucleic acid probes to cyclins, histones, or DNA synthesis enzymes can be used as positive markers for the G1/S transition. For cells that have received the DP gene cassette, stimulation of the cell cycle is manifested in an increased mitotic index, detected by staining for mitotic figures using a DNA dye such as DAPI or Hoechst 33258. Flow cytometric analysis of DP-expressing cells is expected to show that a high percentage of cells have progressed into or through S-phase. Progression through S-phase will be manifested by fewer cells in G1 and/or more rapid cycling times (i.e. shorter G1 stage). A higher percentage of cells are labeled when cell cycle stage-specific probes are used, as mentioned above.

To assess the effect on transgene integration, growth of bialaphos-resistant colonies on selective medium is a reliable assay. Within 1-7 days after DNA introduction (for example, of an expression cassette such as UBI::PAT~GFP::pinII (PAT~GFP fusion) and a Dp-expressing cassette such as UBI::Dp::pinII), the embryos are moved onto culture medium containing 3 mg/l of the selective agent bialaphos. Embryos, and later callus, are transferred to fresh selection plates every 2 weeks. After 6-8 weeks, transformed calli are recovered. Transgenic callus containing the introduced genes can be verified using PCR and Southern analysis. Northern analysis can also be used to verify which calli are expressing the bar gene, and whether the DP gene is being expressed at levels above normal wild-type cells (based on hybridization of probes to freshly isolated mRNA population from the cells). In immature embryos that had transient, elevated DP expression, higher numbers of stable transformants are recovered (likely a direct result of increased integration frequencies). Increased transgene integration frequency can also be assessed using such well-established labeling methods such as in situ hybridization.

For this specific application (using transient DP-mediated cell cycle stimulation to increase transient integration frequencies), it may be desirable to reduce the likelihood of ectopic stable expression of the DP gene. Strategies for transient-only expression can be used. This includes delivery of RNA (transcribed from the DP gene) or DP protein along with the transgene cassettes to be integrated to enhance transgene integration by transient stimulation of cell division. Using well-established methods to produce DP-RNA, this can then be purified and introduced into maize cells using physical methods such as microinjection, bombardment, electroporation or silica fiber methods. For protein delivery, the gene is first expressed in a bacterial or baculoviral system, the protein purified and then introduced into maize cells using physical methods such as microinjection, bombardment, electroporation or silica fiber methods. Alternatively, DP proteins are delivered from *Agrobacterium tumefaciens* into plant cells in the form of fusions to *Agrobacterium* virulence proteins. Fusions are constructed between DP and bacterial virulence proteins such as VirE2, VirD2, or VirF which are known to be delivered directly into plant cells. Fusions are constructed to retain both those properties of bacterial virulence proteins required to mediate delivery into plant cells and the DP activity required for enhancing

transgene integration (WO 99/61619 incorporated by reference herewithin). This method should ensure a high frequency of simultaneous co-delivery of T-DNA and functional DP protein into the same host cell. The methods above represent various means of using the DP gene or its encoded product to transiently stimulate DNA replication and cell division, which in turn enhances transgene integration by providing an improved cellular/molecular environment for this event to occur.

Example 5: Altering DP expression stimulates the cell cycle and growth

Based on results in other eukaryotes, expression of ZmDP genes stimulates the G1/S transition and promotes cell division. This increase in division rate is assessed in a number of different manners, more rapid incorporation of radiolabeled nucleotides, and faster growth (i.e. more biomass accumulation). Delivery of the ZmDP in an appropriate plant expression cassette is accomplished through numerous well-established methods for plant cells, including for example particle bombardment, sonication, PEG treatment or electroporation of protoplasts, electroporation of intact tissue, silica-fiber methods, microinjection or *Agrobacterium*-mediated transformation. The result of ZmDP gene expression will be to stimulate the G1/S transition and hence cell division, providing the optimal cellular environment for integration of introduced genes (as per Example 1). This will trigger a tissue culture response (cell divisions) in genotypes that typically do not respond to conventional culture techniques (recalcitrant), or stimulate growth of transgenic tissue beyond the normal rates observed in wild-type (non-transgenic) tissues. To demonstrate this, the DP gene is cloned into a cassette with a constitutive promoter (i.e. either a strong maize promoter such as the ubiquitin promoter including the first ubiquitin intron, or a weak constitutive promoter such as nos). Either particle-mediated DNA delivery or *Agrobacterium*-mediated delivery are used to introduce the nos::ZmDP::pinII-containing plasmid along with a UBI::*bar*::pinII-containing plasmid into maize cells capable of growth on suitable maize culture medium. Such competent cells can be from maize suspension culture, callus culture on solid medium, freshly isolated immature embryos or meristem cells. A model system is used where immature embryos of the Hi-II genotype are used as the target for co-delivery of these two plasmids, and within 1-7 days the embryos are moved onto culture medium containing 3 mg/l of the selective agent bialaphos. Embryos, and later callus, are transferred to

fresh selection plates every 2 weeks. After 6-8 weeks, transformed calli are recovered. In treatments where both the *bar* gene and DP gene have been transformed into immature embryos, a higher number of growing calli are recovered on the selective medium and callus growth is stimulated (relative to treatments with the *bar* gene alone). When the DP gene is introduced without any additional selective marker, transgenic calli can be identified by their ability to grow more rapidly than surrounding wild-type (non-transformed) tissues. Transgenic callus can be verified using PCR and Southern analysis. Northern analysis can also be used to verify which calli are expressing the bar gene, and which are expressing the maize DP gene at levels above normal wild-type cells (based on hybridization of probes to freshly isolated mRNA population from the cells).

Inducible Expression:

The DP gene can also be cloned into a cassette with an inducible promoter such as the benzenesulfonamide-inducible promoter. The expression vector is co-introduced into plant cells and after selection on bialaphos, the transformed cells are exposed to the safener (inducer). This chemical induction of DP expression should result in stimulated G1/S transition and more rapid cell division. The cells are screened for the presence of ZmDP RNA by northern, or RT-PCR (using transgene specific probes/oligo pairs), for DP-encoded protein using DP-specific antibodies in Westerns or using hybridization. Increased DNA replication is detected using BrdU labeling followed by antibody detection of cells that incorporated this thymidine analogue. Likewise, other cell cycle division assays could be employed, as described above.

Example 6: Control of DP gene expression using tissue-specific or cell-specific promoters provides a differential growth advantage

DP gene expression using tissue-specific or cell-specific promoters stimulates cell cycle progression in the expressing tissues or cells. For example, using a seed-specific promoter will stimulate cell division rate and result in increased seed biomass. Alternatively, driving DP expression with a strongly-expressed, early, tassel-specific promoter will enhance development of this entire reproductive structure. Expression of DP genes in other cell types and/or at different stages of development will similarly stimulate cell division rates. For

example, root-specific or root-preferred expression of DP is expected to result in larger roots and faster growth (i.e. more biomass accumulation).

Example 7. Meristem Transformation

Meristem transformation protocols rely on the transformation of apical initials or cells that can become apical initials following reorganization due to injury or selective pressure. The progenitors of these apical initials differentiate to form the tissues and organs of the mature plant (i.e. leaves, stems, ears, tassels, etc.). The meristems of most angiosperms are layered with each layer having its own set of initials. Normally in the shoot apex these layers rarely mix. In maize the outer layer of the apical meristem, the L1, differentiates to form the epidermis while descendents of cells in the inner layer, the L2, give rise to internal plant parts including the gametes. The initials in each of these layers are defined solely by position and can be replaced by adjacent cells if they are killed or compromised. Meristem transformation frequently targets a subset of the population of apical initials and the resulting plants are chimeric. If for example, 1 of 4 initials in the L1 layer of the meristem are transformed only ¼ of epidermis would be transformed. Selective pressure can be used to enlarge sectors but this selection must be non-lethal since large groups of cells are required for meristem function and survival. Transformation of an apical initial with a DP expression cassette under the expression of a promoter active in the apical meristem (either meristem specific or constitutive) would allow the transformed cells to grow faster and displace wildtype initials driving the meristem towards homogeneity and minimizing the chimeric nature of the plant body. To demonstrate this, the DP gene is cloned into a cassette with a promoter that is active within the meristem (i.e. either a strong constitutive maize promoter such as the ubiquitin promoter including the first ubiquitin intron, or a promoter active in meristematic cells such as the maize histone, cdc2 or actin promoter). Coleoptilar stage embryos are isolated and plated meristem up on a high sucrose maturation medium (see Lowe *et al.*, 1997). The DP expression cassette along with a reporter construct such as Ubi:GUS:pinII can then be co-delivered (preferably 24 hours after isolation) into the exposed apical dome using conventional particle gun transformation protocols. As a control the DP construct can be replaced with an equivalent amount of pUC plasmid DNA. After a week to 10 days of culture on maturation medium the embryos can be

transferred to a low sucrose hormone-free germination medium. Leaves from developing plants can be sacrificed for GUS staining. Transient expression of the DP gene in meristem cells, through stimulation of the G1→S transition, will result in greater integration frequencies and hence more numerous transgenic sectors. Integration and expression of the DP gene will impart a competitive advantage to expressing cells resulting in a progressive enlargement of the transgenic sector. Due to the enhanced growth rate in DP-expressing meristem cells, they will supplant wild-type meristem cells as the plant continues to grow. The result will be both enlargement of transgenic sectors within a given cell layer (i.e. periclinal expansion) and into adjacent cell layers (i.e. anticlinal invasions). As an increasingly large proportion of the meristem is occupied by DP-expressing cells, the frequency of DP germline inheritance should go up accordingly.

Example 8. Use of Flp/Frt system to excise the DP cassette

In cases where the DP gene has been integrated and DP expression is useful in the recovery of maize transgenics, but is ultimately not desired in the final product, the DP expression cassette (or any portion thereof that is flanked by appropriate FRT recombination sequences) can be excised using FLP-mediated recombination (see US Patent 5,929,301 issued July 27, 1999 and incorporated by reference herewithin).

The above examples are provided to illustrate the invention but not to limit its scope. Other variants of the invention will be readily apparent to one of ordinary skill in the art and are encompassed by the appended claims. All publications, patents, and patent applications cited herein are hereby incorporated by reference.

WHAT IS CLAIMED IS

1. A transgenic plant comprising at least one expression cassette comprising at least one DP nucleic acid operably linked to a promoter, wherein the nucleic acid is in sense or antisense orientation.

2. The transgenic plant of claim 1, wherein the plant is corn, soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice, barley or millet.

3. A seed from the transgenic plant of claim 1.

4. The seed of claim 3, wherein the seed is from corn, soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice, barley or millet.

5. A method of modulating the level of DP protein in a plant, comprising:
   a) stably transforming a plant cell with a *DP* polynucleotide operably linked to a promoter, wherein the polynucleotide is in sense or antisense orientation;
   b) growing the plant cell under plant growing conditions to produce a regenerated plant; and
   c) inducing expression of the polynucleotide for a time sufficient to modulate DP protein in the plant.

6. The method of claim 5, wherein the *DP* polynucleotide is encoded by a nucleic acid comprising Seq. ID No. 1.

7. The method of claim 5, wherein the plant is corn, soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice, barley and millet.

8. The method of claim 5, wherein DP protein is increased.

9. The method of claim 5, wherein DP protein is decreased.

10. A method for increasing crop yield comprising modulating DP protein expression by the method of claim 5.

11. A method for providing a positive growth advantage in a plant comprising modulating DP protein expression by the method of claim 5.

12. A method for improving transformation frequencies comprising increasing the number of dividing cells to be transformed by modulating DP protein expression by the method of claim 5.

13. A method for modulating endoreduplication comprising modulating the expression of DP protein in the endosperm of corn, sorghum, wheat, rice, barley or millet by the method of claim 5, wherein the promoter is an endosperm-preferred promoter.

14. A method for identifying DP interacting proteins comprising adducting a DP nucleic acid sequence to a second nucleic acid sequence encoding a DNA-binding domain.

15. A method for transiently modulating the level of DP protein activity in plant cells comprising introducing into the plant cells a member selected from the group consisting of:

a) at least one *DP* polynucleotide, operably linked to a promoter, wherein the polynucleotide is in sense or antisense orientation;

b) at least one polynucleotide encoded by a nucleic acid comprising Seq. ID No. 1, operably linked to a promoter, wherein the polynucleotide is in sense or antisense orientation;

c) at least one *DP* RNA, wherein the RNA is in sense or antisense orientation;

d) at least one DP polynucleotide, wherein the polynucleotide is RNA in sense or antisense orientation;

e) at least one double stranded *DP* RNA, wherein the double-stranded RNA comprises the entire span of the *DP* gene or a portion of the polynucleotide;

f) at least one DP polypeptide;

g) at least one polypeptide encoded by a nucleic acid comprising Seq. ID No. 1;

h) an antibody directed against DP; and

i) an antisense oligonucleotide that complements and binds to its designated target sequence within the *DP* RNA.

16. The method of claim 15 comprising at least one *DP* RNA, wherein the RNA is in sense or antisense orientation.

17. The method of claim 15 comprising at least one double stranded *DP* RNA, wherein the double-stranded RNA comprises the entire span of the *DP* gene or a portion of the polynucleotide.

18. The method of claim 15 comprising an antibody directed against DP.

19. The method of claim 15 comprising at least one polynucleotide encoded by a nucleic acid comprising Seq. ID No. 1, operably linked to a promoter, wherein the polynucleotide is in sense or antisense orientation.

20. The method of claim 19 wherein the promoter is inducible or constitutive.

21. A method for transiently modulating the level of DP protein activity in plant cells comprising introducing into the plant cells at least one *DP* polynucleotide, operably linked to a promoter, wherein the polynucleotide is in sense or antisense orientation.

22. A method for transiently modulating the level of DP protein activity in plant cells comprising introducing into the plant cells at least one *DP* RNA, wherein the RNA is in sense or antisense orientation.

23. A method for improving transformation frequencies comprising increasing the number of dividing cells to be transformed by modulating *DP* expression by the method of claim 15.

24. A method of modulating cell number in a tissue of a plant comprising;

a) stably transforming a plant cell with an *DP* polynucleotide operably linked to a promoter, wherein the polynucleotide is in sense or antisense orientation;

b) growing the plant cell under plant growing conditions to produce a regenerated plant; and

c) inducing expression of the polynucleotide for a time sufficient to modulate DP protein in the plant.

25. The method of claim 24, wherein the tissue is selected from the group consisting of root, seed, tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leaf, stem, seed, trunk, meristem and fruit.

26. A method of transient cell cycle stimulation in a plant cell comprising;

a) introducing at least one DP RNA or DP protein; and

b) introducing at least one transgene cassette to at least one plant cell to stimulate the cell cycle.

27. The method of claim 26 further comprising growing the plant cell under plant growing conditions to produce a regenerated plant.

28. The method of Claim 26 wherein the DP RNA is introduced to the plant cell through microinjection, bombardment, electroporation or silicon fiber methods.

29. The method of Claim 26 wherein the DP protein is introduced to the plant cell by *Agrobacterium tumefaciens*.

30. The method of Claim 26 wherein the plant is recalcitrant.

31. A protein the forms a functional dimer with an E2F polypeptide.

SEQUENCE LISTING

<110> Pioneer Hi-Bred Intl. Inc.

<120> CELL CYCLE NUCLEIC ACIDS, POLYPEPTIDES AND USES THEREOF

<130> 1099-PCT

<160> 6

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 900
<212> DNA
<213> Zea mays

<220>
<221> CDS
<222> (1)...(897)

<221> misc_feature
<222> (1)...(900)
<223> n = A,T,C or G

<400> 1

```
nnn gac agc acg ttc ctc cgc ttg aat aat ctc gac atc aac ggc gac      48
Xaa Asp Ser Thr Phe Leu Arg Leu Asn Asn Leu Asp Ile Asn Gly Asp
 1               5                  10                  15

gac gcg ccg tcg tcg cag gct cct acg agc aag aag aaa agg aga ggc      96
Asp Ala Pro Ser Ser Gln Ala Pro Thr Ser Lys Lys Lys Arg Arg Gly
            20                  25                  30

aca cgg gca gtg ggt cct gat aaa ggt aac cgg gga ctg cgc cag ttt     144
Thr Arg Ala Val Gly Pro Asp Lys Gly Asn Arg Gly Leu Arg Gln Phe
        35                  40                  45

agt atg aaa gtt tgt gag aaa gtt gaa agt aaa ggg aga aca aca tat     192
Ser Met Lys Val Cys Glu Lys Val Glu Ser Lys Gly Arg Thr Thr Tyr
    50                  55                  60

aat gag gtg gca gat gaa ctt gtt gct gag ttt aca gac ccc aac aat     240
Asn Glu Val Ala Asp Glu Leu Val Ala Glu Phe Thr Asp Pro Asn Asn
65                  70                  75                  80

aat att gag gca cca gat cct gat aac cct aac gcg caa caa tat gat     288
Asn Ile Glu Ala Pro Asp Pro Asp Asn Pro Asn Ala Gln Gln Tyr Asp
                85                  90                  95

gag aaa aat ata cga cga aga gtt tat gat gct ttg aat gtt ctg atg     336
Glu Lys Asn Ile Arg Arg Arg Val Tyr Asp Ala Leu Asn Val Leu Met
            100                 105                 110

gct atg gac att ata tct aaa gat aaa aag gag atc cag tgg aag ggc     384
Ala Met Asp Ile Ile Ser Lys Asp Lys Lys Glu Ile Gln Trp Lys Gly
        115                 120                 125

ttg ccg cgt act agt ata agt gac att gaa gaa atg aag acg gag ctt     432
```

```
Leu Pro Arg Thr Ser Ile Ser Asp Ile Glu Glu Met Lys Thr Glu Leu
    130                 135                 140

gtg gga ctg aaa ggt agg att gaa aag aaa agt gct tat cta cag gag      480
Val Gly Leu Lys Gly Arg Ile Glu Lys Lys Ser Ala Tyr Leu Gln Glu
145                 150                 155                 160

cta caa gac caa tat gta ggt ctg caa aac ctg att caa cga aat gaa      528
Leu Gln Asp Gln Tyr Val Gly Leu Gln Asn Leu Ile Gln Arg Asn Glu
                165                 170                 175

cag tca tat ggt tca gga aac acg cct tct ggt gga gtg gct ttg cca      576
Gln Ser Tyr Gly Ser Gly Asn Thr Pro Ser Gly Gly Val Ala Leu Pro
            180                 185                 190

ttt atc cta gtt cag acc cga cct cat gct acc gtg gaa gtt gag ata      624
Phe Ile Leu Val Gln Thr Arg Pro His Ala Thr Val Glu Val Glu Ile
        195                 200                 205

tca gaa gat atg cag ctg gtg cat ttt gac ttc aat agc acc cca ttc      672
Ser Glu Asp Met Gln Leu Val His Phe Asp Phe Asn Ser Thr Pro Phe
    210                 215                 220

gag ctg cac gac gac tca tac gtc cta aaa gaa atg cga ttc tgt gga      720
Glu Leu His Asp Asp Ser Tyr Val Leu Lys Glu Met Arg Phe Cys Gly
225                 230                 235                 240

aga gaa caa cat gac agc act caa gag tcg ata tca aat gga ggt gag      768
Arg Glu Gln His Asp Ser Thr Gln Glu Ser Ile Ser Asn Gly Gly Glu
                245                 250                 255

agc tca agc gtg tca aat att tat tgg caa cac gta cag cat gtg gaa      816
Ser Ser Ser Val Ser Asn Ile Tyr Trp Gln His Val Gln His Val Glu
            260                 265                 270

agg cca aac aat ggc aca ggt agg tta ccg agc tca ccg cct att cca      864
Arg Pro Asn Asn Gly Thr Gly Arg Leu Pro Ser Ser Pro Pro Ile Pro
        275                 280                 285

ggg atc ttg aaa ggg cgt gcg aag cac gag cac taa                      900
Gly Ile Leu Lys Gly Arg Ala Lys His Glu His
    290                 295
```

<210> 2
<211> 299
<212> PRT
<213> Zea mays

<220>
<221> VARIANT
<222> (1)...(299)
<223> Xaa = Any Amino Acid

<400> 2

```
Xaa Asp Ser Thr Phe Leu Arg Leu Asn Asn Leu Asp Ile Asn Gly Asp
 1               5                  10                  15
Asp Ala Pro Ser Ser Gln Ala Pro Thr Ser Lys Lys Lys Arg Arg Gly
            20                  25                  30
Thr Arg Ala Val Gly Pro Asp Lys Gly Asn Arg Gly Leu Arg Gln Phe
        35                  40                  45
Ser Met Lys Val Cys Glu Lys Val Glu Ser Lys Gly Arg Thr Thr Tyr
```

```
     50                  55                  60
Asn Glu Val Ala Asp Glu Leu Val Ala Glu Phe Thr Asp Pro Asn Asn
65                  70                  75                  80
Asn Ile Glu Ala Pro Asp Pro Asp Asn Pro Asn Ala Gln Gln Tyr Asp
                85                  90                  95
Glu Lys Asn Ile Arg Arg Arg Val Tyr Asp Ala Leu Asn Val Leu Met
            100                 105                 110
Ala Met Asp Ile Ile Ser Lys Asp Lys Lys Glu Ile Gln Trp Lys Gly
        115                 120                 125
Leu Pro Arg Thr Ser Ile Ser Asp Ile Glu Glu Met Lys Thr Glu Leu
    130                 135                 140
Val Gly Leu Lys Gly Arg Ile Glu Lys Lys Ser Ala Tyr Leu Gln Glu
145                 150                 155                 160
Leu Gln Asp Gln Tyr Val Gly Leu Gln Asn Leu Ile Gln Arg Asn Glu
                165                 170                 175
Gln Ser Tyr Gly Ser Gly Asn Thr Pro Ser Gly Gly Val Ala Leu Pro
            180                 185                 190
Phe Ile Leu Val Gln Thr Arg Pro His Ala Thr Val Glu Val Glu Ile
        195                 200                 205
Ser Glu Asp Met Gln Leu Val His Phe Asp Phe Asn Ser Thr Pro Phe
    210                 215                 220
Glu Leu His Asp Asp Ser Tyr Val Leu Lys Glu Met Arg Phe Cys Gly
225                 230                 235                 240
Arg Glu Gln His Asp Ser Thr Gln Glu Ser Ile Ser Asn Gly Gly Glu
                245                 250                 255
Ser Ser Ser Val Ser Asn Ile Tyr Trp Gln His Val Gln His Val Glu
            260                 265                 270
Arg Pro Asn Asn Gly Thr Gly Arg Leu Pro Ser Ser Pro Pro Ile Pro
        275                 280                 285
Gly Ile Leu Lys Gly Arg Ala Lys His Glu His
    290                 295
```

<210> 3
<211> 32
<212> DNA
<213> Zea mays

<400> 3
gacagcacgt tcctccgctt gaataatctc ga
32

<210> 4
<211> 33
<212> DNA
<213> Zea mays

<400> 4
ttagtgctcg tgcttcgcac gccctttcaa gat
33

<210> 5
<211> 30
<212> DNA
<213> Zea mays

<400> 5
gacagcacgt tcctccgctt gaataatctc
30

<210> 6
<211> 31
<212> DNA

<213> Zea mays

<400> 6
ttagtgctcg tgcttcgcac gccctttcaa g
31

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 00/03651

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C07K14/415 C12N15/82 C12N15/63 A01H1/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N A01H C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, PAJ, STRAND, BIOSIS, FSTA

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 28334 A (PROLIFIX LTD.) 2 July 1998 (1998-07-02) page 3, line 17 - line 19 page 16, line 31 -page 17, line 2 page 26, line 7 - line 18 ---	14-16, 24,25,31
P,X	WO 99 53075 A (E.I. DU PONT DE NEMOURS AND COMPANY) 21 October 1999 (1999-10-21) page 2, line 28 - line 30 page 8, line 24 - line 38 page 11, line 33 - line 36 --- -/--	1-5,7-9, 15,16, 26-29,31

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
30 June 2000	17/07/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Schönwasser, D

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	MAGAE J ET AL: "NUCLEAR LOCALIZATION OF DP AND E2F TRANSCRIPTION FACTORS BY HETERODIMERIC PARTNERS AND RETINOBLASTOMA PROTEIN FAMILY MEMBERS" JOURNAL OF CELL SCIENCE,GB,ESSEX, vol. 109, no. 7, 1 August 1996 (1996-08-01), pages 1717-1726, XP002041650	1-29
X	the whole document	31
A	SHAN B. ET AL.: "The Molecular Basis of E2F-1/DP-1-induced S-Phase Entry and Apoptosis" CELL GROWTH & DIFFERENTIATION, vol. 7, no. 6, June 1996 (1996-06), pages 689-697, XP000864681	1-29
X	the whole document	31
A	WU C.-L. ET AL.: "Expression of Dominant-Negative Mutant DP-1 Blocks Cell Cycle Progression in G1" MOLECULAR AND CELLULAR BIOLOGY, vol. 16, no. 7, July 1996 (1996-07), pages 3698-3706, XP000915116	26-29
X	the whole document	31

INTERNATIONAL SEARCH REPORT

Information on patent family members

Intern nal Application No
PCT/US 00/03651

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9828334 A	02-07-1998	AU 5330498 A	17-07-1998
		EP 0956299 A	17-11-1999
WO 9953075 A	21-10-1999	AU 3478399 A	01-11-1999
		AU 3478599 A	01-11-1999
		WO 9953069 A	21-10-1999